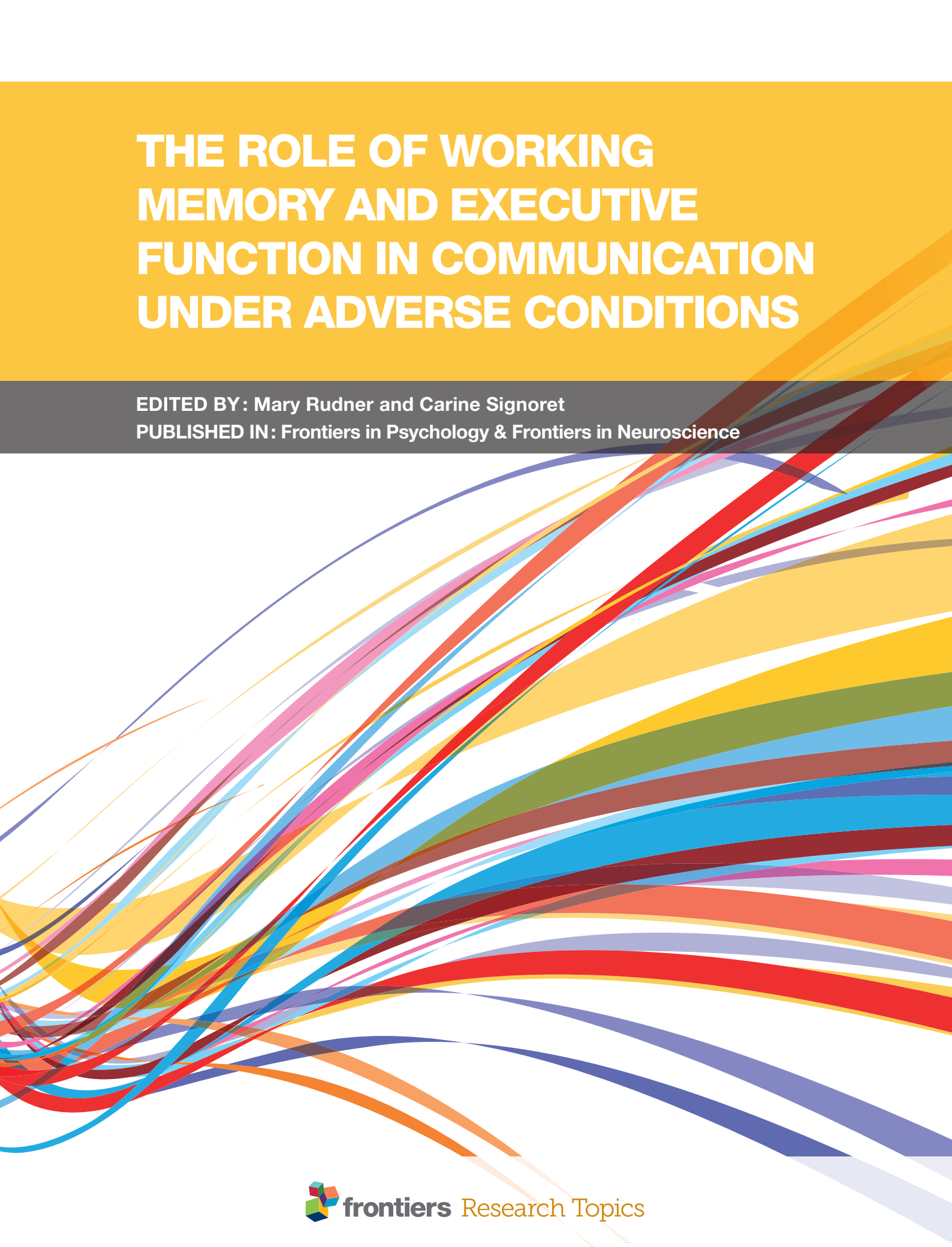
THE ROLE OF WORKING MEMORY AND EXECUTIVE FUNCTION IN COMMUNICATION UNDER ADVERSE CONDITIONS

EDITED BY: Mary Rudner and Carine Signoret
PUBLISHED IN: Frontiers in Psychology & Frontiers in Neuroscience



frontiers Research Topics



frontiers

Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714
ISBN 978-2-88919-861-0
DOI 10.3389/978-2-88919-861-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

THE ROLE OF WORKING MEMORY AND EXECUTIVE FUNCTION IN COMMUNICATION UNDER ADVERSE CONDITIONS

Topic Editors:
Mary Rudner, Linköping University, Sweden
Carine Signoret, Linnaeus Centre for Hearing and Deafness (HEAD), Sweden

Communication is vital for social participation. However, communication often takes place under suboptimal conditions. This makes communication harder and less reliable, leading at worst to social isolation. In order to promote participation, it is necessary to understand the mechanisms underlying communication in different situations. Human communication is often speech based, either oral or written, but may also involve gesture, either accompanying speech or in the form of sign language. For communication to be achieved, a signal generated by one person has to be perceived by another person, attended to, comprehended and responded to. This process may be hindered by adverse conditions including factors that may be internal to the sender (e.g. incomplete or idiosyncratic language production), occur during transmission (e.g. background noise or signal processing) or be internal to the receiver (e.g. poor grasp of the language or sensory impairment). The extent to which these factors interact to generate adverse conditions may differ across the lifespan. Recent work has shown that successful speech communication under adverse conditions is associated with good cognitive capacity including efficient working memory and executive abilities such as updating and inhibition. Further, frontoparietal networks associated with working memory and executive function have been shown to be activated to a greater degree when it is harder to achieve speech comprehension. To date, less work has focused on sign language communication under adverse conditions or the role of gestures accompanying speech communication under adverse conditions. It has been proposed that the role of working memory in communication under such conditions is to keep fragments of an incomplete signal in mind, updating them as appropriate and inhibiting irrelevant information, until an adequate match can be achieved with lexical and semantic representations held in long term memory. Recent models of working memory highlight an episodic buffer whose role is the multimodal integration of information from the senses and long term memory. It is likely that the episodic buffer plays a key role in communication under adverse conditions.

The aim of this research topic is to draw together multiple perspectives on communication under adverse conditions including empirical and theoretical approaches. This will facilitate a scientific exchange among individual scientists and groups studying different aspects of communication under adverse conditions and/or the role of cognition in communication. As such, this topic belongs firmly within the field of Cognitive Hearing Science. Exchange of ideas among scientists

with different perspectives on these issues will allow researchers to identify and highlight the way in which different internal and external factors interact to make communication in different modalities more or less successful across the lifespan. Such exchange is the forerunner of broader dissemination of results which ultimately, may make it possible to take measures to reduce adverse conditions, thus facilitating communication. Such measures might be implemented in relation to the built environment, the design of hearing aids and public awareness.

Citation: Rudner, M., Signoret, C., eds. (2016). The Role of Working Memory and Executive Function in Communication under Adverse Conditions. Lausanne: Frontiers Media.
doi: 10.3389/978-2-88919-861-0

Table of Contents

06 ***Editorial: The Role of Working Memory and Executive Function in Communication under Adverse Conditions***
Mary Rudner and Carine Signoret

09 ***The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests***
Antje Heinrich, Helen Henshaw and Melanie A. Ferguson

23 ***Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss***
Sherri L. Smith and M. Kathleen Pichora-Fuller

38 ***How does susceptibility to proactive interference relate to speech recognition in aided and unaided conditions?***
Rachel J. Ellis and Jerker Rönnberg

45 ***Keeping track of who said what: Performance on a modified auditory n-back task with young and older adults***
Gary R. Kidd and Larry E. Humes

60 ***The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults***
Karen A. Doherty and Jamie L. Desjardins

69 ***Gated auditory speech perception: effects of listening conditions and cognitive capacity***
Shahram Moradi, Björn Lidestam, Amin Saremi and Jerker Rönnberg

82 ***Relatively effortless listening promotes understanding and recall of medical instructions in older adults***
Roberta M. DiDonato and Aimée M. Surprenant

102 ***Hearing loss impacts neural alpha oscillations under adverse listening conditions***
Eline B. Petersen, Malte Wöstmann, Jonas Obleser, Stefan Stenfelt and Thomas Lunner

113 ***Cognitive spare capacity: evaluation data and its association with comprehension of dynamic conversations***
Gitte Keidser, Virginia Best, Katrina Freeston and Alexandra Boyce

127 ***Memory performance on the Auditory Inference Span Test is independent of background noise type for young adults with normal hearing at high speech intelligibility***
Niklas Rönnberg, Mary Rudner, Thomas Lunner and Stefan Stenfelt

138 ***Costs of switching auditory spatial attention in following conversational turn-taking***
Gaven Lin and Simon Carlile

149 ***Working memory and intelligibility of hearing-aid processed speech***
Pamela E. Souza, Kathryn H. Arehart, Jing Shen, Melinda Anderson and James M. Kates

163 ***Relating hearing loss and executive functions to hearing aid users' preference for, and speech recognition with, different combinations of binaural noise reduction and microphone directionality***
Tobias Neher

177 ***Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss***
Melanie A. Ferguson and Helen Henshaw

184 ***Intrinsic and extrinsic motivation is associated with computer-based auditory training uptake, engagement, and adherence for people with hearing loss***
Helen Henshaw, Abby McCormack and Melanie A. Ferguson

197 ***Subjective ratings of masker disturbance during the perception of native and non-native speech***
Lisa Kilman, Adriana A. Zekveld, Mathias Hällgren and Jerker Rönnberg

204 ***Speech intelligibility and recall of first and second language words heard at different signal-to-noise ratios***
Staffan Hygge, Anders Kjellberg and Anatole Nöstl

211 ***The contribution of phonological knowledge, memory, and language background to reading comprehension in deaf populations***
Elizabeth A. Hirshorn, Matthew W. G. Dye, Peter Hauser, Ted R. Supalla and Daphne Bavelier

227 ***Deaf children's non-verbal working memory is impacted by their language experience***
Chloë Marshall, Anna Jones, Tanya Denmark, Kathryn Mason, Joanna Atkinson, Nicola Botting and Gary Morgan

239 ***Executive functions in mono-and bilingual children with language impairment – issues for speech-language pathology***
Olof Sandgren and Ketty Holmström

244 ***On the interaction of speakers' voice quality, ambient noise and task complexity with children's listening comprehension and cognition***
Viveka Lyberg-Åhlander, K. J. Brännström and Birgitta S. Sahlén

249 ***Working memory and referential communication—multimodal aspects of interaction between children with sensorineural hearing impairment and normal hearing peers***
Olof Sandgren, Kristina Hansson and Birgitta Sahlén

253 ***Oral communication in individuals with hearing impairment—considerations regarding attentional, cognitive and social resources***
Ulrike Lemke and Sigrid Scherpiet

260 ***Cognitive aging and hearing acuity: modeling spoken language comprehension***
Arthur Wingfield, Nicole M. Amichetti and Amanda Lash

frontiers in Psychology

EDITORIAL
published: 11 February 2016
doi: 10.3389/fpsyg.2016.00148

Editorial: The Role of Working Memory and Executive Function in Communication under Adverse Conditions

Mary Rudner and Carine Signoret*

Department of Behavioural Sciences and Learning, Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

Keywords: cognition, communication, adverse conditions, hearing, deafness

The Editorial on the Research Topic

The Role of Working Memory and Executive Function in Communication under Adverse Conditions

Communication is fundamental for social participation with communication difficulties often leading to social isolation and depression. Nevertheless, everyday communication is often hindered either by internal factors such as sensory loss, or by external factors including the background noise that commonly occurs in places where people meet, such as restaurants, schools, and railway stations. In such adverse conditions, working memory and executive functions have been proposed to play a critical role in communication. Thus, the role of cognition in hearing is a central theme in the field of Cognitive Hearing Science and has crystalized as one of the main themes of this research topic. This is reflected in papers reporting the role of cognition in hearing in persons with varying sensory and cognitive status and varying degrees of language knowledge, over the lifespan. Another theme represented in this topic is rehabilitation in the form of amplification and training. Importantly, the broad remit of the research topic is reflected in papers addressing cognition and communication in children with sensory and cognitive issues as well as adults and children who are profoundly deaf and use sign language to communicate. Apart from the impressive number of empirical studies, there are several theoretical contributions to the field.

The observation of consistent correlations between cognitive skills and the ability to understand speech under adverse conditions has played an important role in driving the field of Cognitive Hearing Science. In particular, it has been reported repeatedly that working memory explains variance in the ability to recognize speech in noise above and beyond differences in hearing thresholds. In the current research topic, Heinrich et al. report a study showing, in line with previous work, that individual differences in sensory and cognitive skills explain variance in the ability of older listeners with mild sensorineural hearing loss to process speech. However, they also show that the relative explanatory value of these skills depends on the linguistic demands of the particular speech test, with hearing sensitivity being more important at the phoneme level and cognition at the sentence level. Further, they reported associations between self-reported aspects of auditory functioning and speech intelligibility. Smith and Pichora-Fuller compared performance on the reading span test (RS), which is a well-established measure of working memory delivered visually, and the Word Auditory Recognition and Recall Measure (WARRM), a newer measure of working memory with auditory delivery, which they propose is more ecologically valid. WARRM performance was better and more varied than RS performance in all groups tested (young adults with normal hearing, young-older hard-of-hearing adults and old-older hard-of-hearing adults) and the authors suggested that this pattern of performance indicates WARRM

OPEN ACCESS

Edited and reviewed by:
Isabelle Peretz,
Université de Montréal, Canada

****Correspondence:***
Mary Rudner
mary.rudner@liu.se

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *24 December 2015*
Accepted: *27 January 2016*
Published: *11 February 2016*

Citation:
Rudner M and Signoret C (2016) Editorial: The Role of Working Memory and Executive Function in Communication under Adverse Conditions. Front. Psychol. 7:148. doi: 10.3389/fpsyg.2016.00148

may be a useful clinical test. However, no consistent pattern of correlations was found between the two cognitive measures and measures of the ability to understand speech in noise. Smith and Pichora-Fuller suggest that there is a need for a more consistent approach to determine in more ecologically relevant conditions associations between working memory and speech understanding. During speech comprehension, encoding of new memories may be hampered by interference from established memories; this is known as proactive interference. Ellis and Rönnberg studied whether the ability to suppress such interference was associated with speech recognition in noise in older hard-of-hearing adults. In line with previous work on individuals with normal hearing, they did find an association, but only when hearing was unaided. They suggested that the cognitive flexibility reflected by performance on their cognitive task is a key factor in listening ability.

Experimental approaches are adopted by another set of studies studying the role of cognition in communication. Kidd and Humes used an auditory working memory task to determine differences in the ability of older and younger listeners to keep track of who said what. They found that although older listeners were slower, they were almost as accurate as younger listeners. However, older listeners did not benefit from consistent mapping of target speaker and location in the same way as younger listeners. Doherty and Desjardins investigated how amplification influenced auditory working memory performance in hard-of-hearing listeners who were fitted with hearing aids for the first time. They found that amplification improved working memory and the overall pattern of results suggested that this was due to perceptual benefit rather than cognitive change. Moradi et al. used a gating paradigm to determine whether background noise influences how much of the auditory signal is needed before identification of its linguistic content is achieved. Results showed that more auditory signal was required in noise and that this effect was modulated by both working memory and executive function. DiDonato and Surprenant investigated how speech manner influences the ability of older and younger listeners to remember auditory information with ecological relevance. They found that older listeners could remember medical information better when it was presented clearly rather than conversationally, even in background noise. The electrophysiology study by Petersen et al. investigated how working memory indexed by neural oscillations in a low frequency (alpha) band, is influenced by increasing stimulus degradation and working memory load, in hard-of-hearing individuals. In line with previous work in individuals with normal hearing thresholds, performance decreased and alpha power increased with greater stimulus degradation and working memory load. However, at the highest levels of degradation and working memory load, alpha power dropped for the participants with the greatest degree of hearing loss, suggesting a breakdown in an important neural mechanism that may support listening in noise.

If cognitive resources are consumed during listening in noise as indicated by the association between working memory and listening performance, fewer resources, or less cognitive spare capacity (CSC) will be available for higher level processing of the message. The research topic includes a set of studies investigating this phenomenon. In line with previous work, Keidser et al. found that performance on the CSC Test (CSCT) was influenced by some of the manipulated parameters (but not seeing the talker's face) and that there was no consistent relation between CSCT and RS. Further, there was no relation between CSCT and a novel speech comprehension test presented in noise. Using the Auditory Inference Span Test (AIST), a sentence-based test which involves storage and processing of the message, Rönnberg et al. showed that, even when audibility is relatively well-maintained, processing of a spoken message becomes harder for listeners with normal hearing thresholds when noise level increases, but only when the noise is speech-like. This suggests that speech-like background reduces CSC. Lin and Carlile used a version of AIST to investigate the listening costs associated with shifts in spatial attention during conversational turn-taking in listeners with normal hearing thresholds. They found that listening costs were dependent on load and cognitive complexity but not on the nature of the spatial shift.

Hearing aid signal processing is designed to improve speech understanding. It is important to determine whether this is actually the case and at the same time identify any contingent cost in terms of cognitive function. In this research topic, Souza et al. investigated the role of working memory in speech intelligibility in noise with hearing aid signal processing. The data corroborated previous results showing that individuals with low working memory capacity may benefit more from signal processing that better retains the signal envelope. Neher studied whether working memory and executive function were related to speech recognition in noise performance with hearing aid signal processing as well as preference for different hearing aid fittings in older hearing aid users. His study found that working memory was related to performance with directional microphones while executive function was related to preference for noise reduction.

Ferguson and Henshaw reviewed three auditory training studies and conclude that training which combines auditory and cognitive demands is most likely to benefit hard-of-hearing adults in real-world listening situations. Henshaw et al. argue that training benefit is dependent on uptake, engagement and adherence. Their study showed that uptake was associated with extrinsic motivation (e.g., hearing difficulties) while engagement and adherence were influenced by both intrinsic (e.g., a desire to achieve higher scores), and extrinsic (e.g., to help others with hearing loss) motivations.

An atypical language model can lead to particular involvement of working memory and executive function in language processing. Kilman et al. studied the amount of disturbance perceived by hard-of-hearing listeners and listeners with normal hearing thresholds when attending to a target talker against a multitalker background. Speech was either native or non-native. Results showed that hard-of-hearing participants were particularly disturbed by native speech masked by native babble. Hygge et al. investigated how nativeness of speech influenced the ability to recognize and recall speech in different levels of background noise. They found that recall was more sensitive than recognition to both factors and thus a better indicator for the acoustics of learning.

Because profoundly deaf individuals do not have access to sound, reading and other academic skills may develop differently from those of individuals with normal hearing. Hirshorn et al. assessed the impact of language experience on predictors of reading comprehension in deaf readers. They found that while English phonological knowledge best predicted reading comprehension in oral deaf individuals, free recall was a better predictor in deaf native signers. Marshall et al. investigated the relationship between working memory and language in deaf signing children who were either native or non-native users of British Sign Language compared to hearing children. The non-native signers performed less accurately than both the native signers and the hearing children. Further, vocabulary predicted working memory, suggesting that the good language skills resulting from early acquisition are important for development of working memory.

A number of the papers in the research topic report studies investigating cognitive aspects of language development in children with disabilities. In a perspective article, Sandgren and Holmström discuss the clinical challenge of assessing language impairment in bilingual children and present work suggesting that measuring executive function may be a useful approach. In a mini-review Lyberg-Åhlander et al. discuss their recent work investigating how children's listening comprehension is influenced by speaker voice quality and background noise, as well as the child's own cognitive capacity. They highlight risk of underachievement when speech is delivered in a dysphonic (hoarse) voice, especially when the task is simple or the child's capacity is stretched. In another mini-review, Sandgren et al. summarize their work on referential communication showing that while children with sensorineural hearing impairment are active and competent conversational partners, their conversational strategies are distinct from those of their peers with normal hearing, even when the listening situation is optimized.

Finally, two perspective articles round off the research topic. Lemke and Scherpiet discuss communication from an aging perspective, and the psycho-social impact of sensory and cognitive decline. Wingfield et al. discuss the Ease of Language Understanding (ELU) model (Rönnberg et al.) as one of the few attempts to offer a fully encompassing framework for language understanding. They identify its strengths and point out avenues for future work.

Altogether, the articles in this research topic demonstrate the crucial role of cognition, including working memory and executive functions but also cognitive flexibility and cognitive load, in communication under adverse conditions, in different modalities, and over the lifespan.

AUTHOR CONTRIBUTIONS

MR prepared the first draft of the editorial, and CS and MR contributed equally to completion of the final version.

FUNDING

This work was supported by funding from the Swedish Research Council to the Linnaeus Centre HEAD, Swedish Institute for Disability Research, Department of Behavioral Sciences and Learning, Linköping University.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Rudner and Signoret. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 16 June 2015
doi: 10.3389/fpsyg.2015.00782

The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests

Antje Heinrich[1], Helen Henshaw[2] and Melanie A. Ferguson[2,3]*

[1] Medical Research Council Institute of Hearing Research, Nottingham, UK, [2] National Institute for Health Research–Nottingham Hearing Biomedical Research Unit, Otology and Hearing Group, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, UK, [3] Nottingham University Hospitals NHS Trust, Nottingham, UK

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Adriana A. Zekveld,
VU University Medical Center, Netherlands
Piers Dawes,
The University of Manchester, UK

****Correspondence:***
Antje Heinrich,
Medical Research Council Institute of Hearing Research, University Park, Nottingham, NG7 2RD, UK
antje.heinrich@ihr.mrc.ac.uk

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *28 February 2015*
Accepted: *26 May 2015*
Published: *16 June 2015*

Citation:
Heinrich A, Henshaw H and Ferguson MA (2015) The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. Front. Psychol. 6:782. doi: 10.3389/fpsyg.2015.00782

Listeners vary in their ability to understand speech in noisy environments. Hearing sensitivity, as measured by pure-tone audiometry, can only partly explain these results, and cognition has emerged as another key concept. Although cognition relates to speech perception, the exact nature of the relationship remains to be fully understood. This study investigates how different aspects of cognition, particularly working memory and attention, relate to speech intelligibility for various tests. Perceptual accuracy of speech perception represents just one aspect of functioning in a listening environment. Activity and participation limits imposed by hearing loss, in addition to the demands of a listening environment, are also important and may be better captured by self-report questionnaires. Understanding how speech perception relates to self-reported aspects of listening forms the second focus of the study. Forty-four listeners aged between 50 and 74 years with mild sensorineural hearing loss were tested on speech perception tests differing in complexity from low (phoneme discrimination in quiet), to medium (digit triplet perception in speech-shaped noise) to high (sentence perception in modulated noise); cognitive tests of attention, memory, and non-verbal intelligence quotient; and self-report questionnaires of general health-related and hearing-specific quality of life. Hearing sensitivity and cognition related to intelligibility differently depending on the speech test: neither was important for phoneme discrimination, hearing sensitivity alone was important for digit triplet perception, and hearing and cognition together played a role in sentence perception. Self-reported aspects of auditory functioning were correlated with speech intelligibility to different degrees, with digit triplets in noise showing the richest pattern. The results suggest that intelligibility tests can vary in their auditory and cognitive demands and their sensitivity to the challenges that auditory environments pose on functioning.

Keywords: speech perception, cognition, self-report, communication, health-related quality of life, non-verbal intelligence

Introduction

One of the overarching aims of audiological (re)habilitation is to improve communication skills and participation in everyday life by reducing activity limitations and participation restrictions (e.g., Boothroyd, 2007) The success of any intervention, such as hearing aid fitting, can be assessed using different aspects of communication such as behavioral measures of speech perception, or subjective questionnaires of self-reported hearing-related, or generic health-related quality of life (HRQoL). One way of conceptualizing communication and how to measure it, is by placing it within the World Health Organization's International Classification of Functioning, Disability and Health (ICF: WHO, 2001). The ICF framework suggests that an individual's level of functioning is not simply the consequence of an underlying health condition but instead should be thought of as a multifactorial concept that includes a person's body *functions and structures*, the *activities* they perform and the social situations they *participate* in. All of these factors can be subject to external environmental and internal personal influences (Stucki and Grimby, 2004). Conceptualizing hearing, listening, and communication within this framework places hearing loss as a body *function*, listening (e.g., to speech in noise) as *activity*, and communication as *participation* (e.g., Saunders et al., 2005; Hickson and Scarinci, 2007; Danermark et al., 2010). Experimentally it has been shown that while hearing sensitivity affects listening in a variety of situations (Humes and Roberts, 1990; van Rooij and Plomp, 1990) it has also become increasingly clear that hearing loss alone cannot account for speech perception difficulties, particularly in noise (Schneider and Pichora-Fuller, 2000; Wingfield and Tun, 2007). As a consequence, the role of cognition for speech perception has come under scrutiny. Research so far has led to the general agreement that a relationship between cognition and speech-in-noise (SiN) perception exists but the nature and extent of the relationship is less clear. No single cognitive component has emerged as being important for all listening situations, although working memory (WM), specifically as tested by reading span, appears to be important in many situations (for a review, see Akeroyd, 2008).

Crucially, WM has no universally accepted definition. One definition that is widely used particularly in connection with speech perception, posits that WM capacity refers to the ability to simultaneously store and process task-relevant information (Daneman and Carpenter, 1980). Tasks have been designed that differ in the emphasis they put on storage and processing components. An example of a task with an emphasis on the storage component is the Digit Span forward task (Wechsler, 1997), an example of a task that maximizes the processing component is the Reading Span task (Daneman and Carpenter, 1980). Tasks that put a more equal emphasis on both storage and processing aspects are the Digit Span backward and the visual letter monitoring (VLM) task. WM is often correlated with speech perception, particularly when the speech is presented in multi-talker or fluctuating noise. Moreover, this correlation is often larger when the WM task contains a large processing component (Akeroyd, 2008). However, despite these general trends results have been less clear-cut. For instance, some (Desjardins and Doherty, 2013) but not all (Koelewijn et al., 2012) studies showed the expected significant correlation between reading span and SiN perception. In addition, some studies showed significant correlations between SiN perception and forward and backward digit span (Humes et al., 2006), and VLM (Rudner et al., 2008) even though these WM tasks do not maximize the processing component.

Defining WM in terms of storage and processing capability is not the only option. Other definitions of WM emphasize the role of inhibition of irrelevant information (Engle and Kane, 2003), resource-sharing, the ability to divide and switch attention (Barrouillet et al., 2004), and memory updating (Miyake et al., 2000). Importantly, these have also been linked to SiN perception (e.g., Schneider et al., 2010; Mattys et al., 2012). Finally, it is important to note that the recent focus on cognitive contributions does not imply that hearing sensitivity is not important. An approach that considers the interactive effect of both like the current study is most likely to advance our understanding of speech in noise difficulties (Humes et al., 2013).

Another factor that adds complexity to the relationship between speech perception and cognition is the type of speech perception test used. Two aspects important in a speech perception test are the complexity of the target speech and the complexity of the background noise. The target speech can vary from single phonemes to single words to complex sentences, while the background noise can vary from a quiet background to steady-state noise to a highly modulated and linguistically meaningful multi-talker babble. As a result, the same cognitive test can correlate significantly with speech perception when using a more complex sentence perception test (Desjardins and Doherty, 2013; Moradi et al., 2014) but not when using less complex syllables (Kempe et al., 2012). Similarly, correlations with cognitive processes are greater when listening to speech in adverse noisy conditions than when listening in quiet (e.g., van Rooij and Plomp, 1990; Wingfield et al., 2005; Rönnberg et al., 2010). In order to cover a wide range of listening situations with relatively few speech perception tests we varied the complexity of both the target and background signal simultaneously. In the low complexity condition listeners were required to discriminate phonemes in quiet, in the medium condition to recognize words in a steady-state background noise and in the most complex condition to comprehend sentences presented in a modulated noise.

When speech perception is measured in noise, the signal-to-noise ratio (SNR) can be manipulated in one of two ways. First, the noise level is fixed and the signal level of the target is varied, or second, the level of the target is fixed and the level of the noise varied. Both methods of setting SNR are used in speech research (Mayo et al., 1997; Smits et al., 2004, 2013; Vlaming et al., 2011), usually without any discussion on how this methodological variation may affect speech perception. Conversely, in audiology practice, the preferred method for changing SNR is to fix the noise and decrease the signal levels (Wilson et al., 2007), because there is an understanding that increasing the noise level can add a quality of annoyance to the signal that is unrelated to

intelligibility (Nabelek et al., 1991). Using the Digit Triplet Test, we explored the consequences of both methods of adjusting the SNR for speech perception and their relationships with cognitive function and self-report measures.

Self-report questionnaires assess subjective experience. A recent systematic review identified 51 different questionnaires that were used by studies that met the review's specific research requirements (Granberg et al., 2014). Questionnaires can be considered as assessing either generic HRQoL or disease-specific (e.g., hearing) aspects (Chisolm et al., 2007). One example of a generic and widely used HRQoL questionnaire is the EQ-5D (The EuroQol Group, 1990). It assesses an individual's ability to perform activities and measures the resulting limits on levels of participation. However, it has been shown to be insensitive to hearing loss (Chisolm et al., 2007; Grutters et al., 2007). Therefore, an additional set of questions based on the same assessment principles have been developed that extends the EQ-5D and is sensitive to hearing-specific health states such as communication, self-confidence, and family activities (Arlinger et al., 2008). Alternatively, hearing-specific questionnaires can measure activity limitations and participation restrictions, with different questionnaires assessing different aspects of listening. For example, the Auditory Lifestyle and Demand Questionnaire (ALDQ; Gatehouse et al., 1999) assesses listening situations and demands in terms of frequency and importance, the Speech, Spatial, and Qualities of Hearing Questionnaire (SSQ; Gatehouse and Noble, 2004) assesses the listener's ability to perform in particular listening situations, and the Glasgow Hearing Aid Benefit Profile (GHABP; Gatehouse, 1999) assesses activity limitations and participation restrictions associated with listening to speech. However, relatively little is understood about the relationship between different listening situations as measured by hearing-specific questionnaires and performance on various speech perception tests (Cox and Alexander, 1992; Humes et al., 2001).

In addition to examining the relationship between self-report and speech perception in general, we also investigated whether the procedural differences for varying SNRs affect the relationship between speech perception and self-report scores. If for instance setting the SNR by changing the level of noise rather than the signal leads to increased noise levels (as would occur if the SNR for 50% performance threshold is negative), the resulting SNR may become uniquely associated with self-report scales on auditory functioning in noisy environments.

In summary, the current study aimed to assess the relationship between (1) speech perception and cognition, and (2) speech perception and self-report, and how these relationships changed when speech perception tests differed in complexity. Based on previous research we made the following predictions:

Aim 1: Assessing the relationship between speech perception and cognition

(1.1) Speech perception performance will be associated with cognition, and this will be moderated by hearing sensitivity.
(1.2) The contribution of cognition will increase as the complexity of the speech perception task increases.
(1.3) Where procedural differences in identifying SNR occur while the speech and background signals are identical, we expect comparable associations with cognition if these associations are driven by signal complexities and not procedural differences.

Aim 2: Assessing the relationship between speech perception performance and self-reported outcomes

(2.1) Hearing-specific questionnaires will demonstrate a greater association with speech perception performance than generic health measures.
(2.2) Correlations with speech perception performance will be largest for questionnaires that capture aspects of listening important for that particular speech perception test.
(2.3) Procedural differences in identifying SNR for speech perception performance may lead to different associations with self-report scales. In particular, increasing the level of background noise to reduce perceptual accuracy may be uniquely associated with functioning in challenging auditory environments.

By better understanding the relationship between behavioral and subjective measures of listening, this study aims to enable healthcare practitioners and researchers to be more informed in their choice of the outcome measures (either speech perception tests or questionnaires) that relates explicitly to the needs and goals of a particular individual (Gatehouse, 2003) or research question.

Materials and Methods

The data were a subset of a randomized controlled trial to assess the benefits of a home-delivered auditory training program (Ferguson et al., 2014) in which 44 adults with mild sensorineural hearing loss (SNHL) completed outcome measures of speech perception, cognition, and self-report of health and hearing ability. Here, we only examine the baseline data from the participants' initial visit. The study was approved by the Nottingham Research Ethics Committee and Nottingham University Hospitals NHS Trust Research and Development. Signed, informed consent was obtained.

Participants

Participants (29 male, 15 female) were aged 50–74 years old (mean = 65.3 years, SD = 5.7 years) with mild, symmetrical SNHL (mean hearing thresholds averaged across 0.5, 1, 2, and 4 kHz = 32.5 dB HL, SD = 6.0 dB HL, with a left–right difference of <15 dB). All participants spoke English as their first language, and were paid a nominal attendance fee and travel expenses for their visit.

Procedure

Audiometric measurements (middle-ear function and pure-tone air-conduction thresholds) were obtained in a sound-attenuated booth. All other testing (cognitive tests, speech perception tests and self-report questionnaires) took place in a purpose-designed

quiet test room. Outcome measures were administered in the same order for all participants.

Outcome Measures

Audiological

Outer and middle ear functions were checked by otoscopy and standard clinical tympanometry using a GSI Tympstar (Grason-Stadler, Eden Prairie, MN, USA). *Pure-tone air conduction thresholds* (0.25, 0.5, 1, 2, 3, 4, 8 kHz) were obtained for each ear, following the procedure recommended by the British Society of Audiology (British Society of Audiology, 2011), using a Siemens (Crawley, West Sussex, UK) Unity PC audiometer, Sennheiser (Hannover, Germany) HDA-200 headphones, and a B71 Radioear (New Eagle, PA, USA) transducer in a sound-attenuating booth. The better-ear-average (BEA) across octave frequencies 0.5–4 kHz was derived and is reported here.

Cognitive

The *Matrix Reasoning* subtest of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) estimated the non-verbal intelligence quotient (NVIQ). The *Digit Span* (forward, then backward) from the Wechsler Adult Intelligence Scale (WAIS) Third Edition (Wechsler, 1997) was used to estimate auditory WM capacity. Pairs of pre-recorded spoken digit (0–9) sequences were presented at 70 dBA via Sennheiser HD-25 headphones. On successful recall, the sequence was increased by one digit. The test was discontinued when both sequence pairs were incorrectly recalled.

The Visual Letter Monitoring test (VLM) assessed visual WM (Gatehouse, 2003). Ten consonant-vowel-consonant (CVC) words were embedded within an 80-letter sequence displayed sequentially on a computer screen. Participants pressed the keyboard space bar when three consecutive letters formed a recognized CVC word (e.g., M-A-T). The test consisted of two runs, initially with a presentation rate of one letter/2 s, followed by one letter/1 s. Here, only responses to the faster presentation sequence were analyzed in terms of hits (accuracy in %) and reaction time (processing speed in ms).

Two subtests of the *Test of Everyday Attention* (TEA; Robertson et al., 1994) assessed focused and divided attention. In the Telephone Search (Subtest 6, focused attention) participants had to identify 20 designated key symbols, as fast as possible, and ignore all other symbols while searching entries in a simulated classified telephone directory. The score was calculated as a quotient between the total time taken to complete the test divided by the number of symbols detected. The maximum number was 20 and lower values represent superior performance. Divided attention was measured with the Telephone Search (Subtest 7, dual task) that was identical to subtest 6 except that participants had to count a string of 1-kHz tones while searching the directory. The task score was considered separately, and in conjunction with subtext 6 (dual task decrement, DTD). For statistical analyses the scales for both tests were reversed to harmonize the direction of scoring for all cognitive tests with higher scores indicating a better performance in all instances.

Speech Perception

The *Phoneme Discrimination* test measured the discrimination threshold for one phoneme continuum (/a/ to /e/) with 96 steps. Stimuli were delivered through Sennheiser HD-25 headphones at a fixed level of 75 dBA. A three-interval, three-alternative forced choice, oddball paradigm using a step size of 2 combined with a three-down, one-up staircase procedure starting with the second reversal was used to determine the 79% correct point on the psychometric function (Levitt, 1971). Feedback was given. Phoneme discrimination threshold (PD; %) was the average of the last two reversals over 30 trials.

The *Digit Triplet Test* (Smits et al., 2004; Smits and Houtgast, 2005) presented series of three digits against a steady, speech-shaped background noise. Six lists of digits were randomized to minimize order effects. The 50% threshold for digits perception was determined in two ways, (i) the speech level was fixed at 65 dB SPL and the noise level was adaptively varied (DTT_{VN}) (ii) the noise level was fixed at 65 dB SPL and the speech level was adaptively varied (DTT_{VS}). Both noise and speech varied in 2 dB steps in a one-down, one-up paradigm for 27 trials starting with a SNR of +5 dB.

The *Adaptive Sentence List* (ASL; MacLeod and Summerfield, 1990) comprised 30 sentences presented in a 8-Hz modulated noise. Sentences consisted of five words, including three key words (e.g., The lunch was very early), which all needed to be correctly repeated for a sentence to be scored as correct. In keeping with current audiological practice, the noise level was fixed at 60 dBA, and the speech level was adaptively varied first in 10 and 5 dB steps in a one-up, one-down procedure for the first two reversals changing to a three-down, one-up paradigm, and a 2.5 dB step size starting with a SNR of +20 dB. The speech reception threshold was the average SNR of the last two reversals.

All speech perception in noise tests were presented in free-field at a distance of 1 meter. In all speech perception tests a lower score indicates a better performance.

Self-Report of Health-Related Quality of Life (HRQoL)

EQ-5D (The EuroQol Group, 1990) is a standardized generic self-report questionnaire measuring HRQoL. It comprises five questions, each on a three-point scale (no problems, some problems, extreme problems) that assess general life quality as it relates to mobility, self-care, usual activities, pain/discomfort, and affective disorders (depression/anxiety; general EQ-5D). In addition, a set of questions focusing on hearing-specific health states (hearing-specific EQ-5D) was used to assess aspects of life directly related to hearing loss, such as communication, confidence, family activities, social and work activities, and energy level (Arlinger et al., 2008).

Self-Report of Hearing

The ALDQ (Gatehouse et al., 1999) measures frequency and impact of hearing loss by inquiring about a variety of listening situations ($n = 24$). Both dimensions are evaluated on a three-point scale (Frequency: very rarely/sometimes/often; Importance: very little/some importance/very important).

Questions range from listening to sounds of various intensities, to listening to distorted or masked speech to listening to various sound types. Here, an average of both subscales is used where a higher value indicates a richer auditory environment of higher importance to the listener.

The GHABP (Gatehouse, 1999) assesses activity limitations and participation restrictions using four predefined situations (e.g., TV level set to suit other people, conversation with one other person in no background noise, in a busy street, with several people in a group) on a five-point scale (1 = no difficulty to 5 = cannot manage at all). The mean scores for the two subscales of activity limitations and participation restrictions were converted to a percentage and then averaged for an overall score of communication ability.

The SSQ (Gatehouse and Noble, 2004) assesses abilities and experiences of hearing in difficult listening situations. It comprises 49 questions across a variety of hearing domains such as speech perception in a variety of competing contexts (Speech, n = 14), using directional, distance, and movement components to hear (Spatial, $n = 17$) and judging quality of hearing regarding clarity and ability to identify different speakers, musical pieces/instruments, and everyday sounds (Qualities, $n = 18$). Participants rate their hearing ability along a 0–10 visual analog scale for each questions (0 = not at all to 10 = perfectly). Mean scores for each subscale were calculated and averaged for an overall mean score.

Scales were reversed for all further analyses for the general EQ-5D, the hearing-specific EQ-5D, and GHABP in order to assign the highest values to scores of best functioning and richest environment.

Data Analysis

Relationship with Cognitive Tests

Simple Pearson product-moment correlations between each of the four speech perception tests and age, BEA hearing thresholds, and cognitive measures were calculated. Because performance on all but one (phoneme discrimination) speech perception test was significantly correlated with hearing thresholds, partial correlations between speech perception and cognition were calculated by controlling for BEA. Differences in correlations between cognitive tests and speech perception tests were assessed by computing z-values for differences between correlations following Steiger (1980).

A main interest of the study was the predictive value of performance on cognitive tests for each speech perception test. However, the number of cognitive tests was fairly large (seven) for a relatively modest sample size of 44 participants. In order to reduce the number of cognitive tests (predictors) for the subsequent regression analysis, a principal component analysis (PCA) was performed in one of two ways. First, a single component solution, explaining the maximum amount of variance among all seven cognitive tests, was extracted. Second, using an orthogonal rotation with Kaiser Normalization, all components following the Kaiser criterion (KMO) of eigenvalues > 1 were extracted, which in this case resulted in a two-factor solution. Both solutions, the single-factor and the two-factor solution, were subsequently used as predictors in separate two-step forward hierarchical regression analyses in which BEA was always entered in a first step to control for hearing, and the extracted one- or two-factor solutions second. Finally, the influence of hearing and cognition for each of the speech perception tests was simultaneously compared in a canonical correlation analysis (CCA) using multivariate ANOVAs to assess whether the pattern of influence of hearing and cognition differed between the four speech perception tests.

A similar analysis plan was followed for self-report, except for the following two deviations. First, no partial correlations with the control of BEA were computed for self-report measures because hearing loss is an essential component of hearing questionnaires. Second, no principal component solutions were extracted and no regression analyses were performed as self-report measures were not conceptualized as predictors for speech perception performance.

Results

A description of all variables is presented in **Table 1**.

Aim 1: Assessing the Relationship between Speech Perception and Cognition

Prediction 1.1. Speech Perception Performance will be Associated with Cognition, and this will be Moderated by Hearing Sensitivity

Correlational analyses

All Pearson product-moment correlations between speech perception tests, hearing thresholds and cognitive variables that were significant at $p < 0.05$ (two-tailed) are shown as scatter plots in **Figure 1**. All speech perception tests except phoneme discrimination were positively correlated with BEA. Because speech perception performance was measured in SNR for a fixed intelligibility level, a lower SNR translated to better performance. The positive correlation with BEA indicated that better hearing sensitivity was associated with lower SNR values. In addition, sentence perception was negatively correlated with Digit Span backward and focused attention (TEA6) indicating that higher scores on these tasks were associated with better intelligibility. A marginal negative correlation was observed between ASL and dual attention (TEA7) indicating that better ability to divide attention was associated with better intelligibility and as a result a lower SNR. DTT_{VS} was marginally positively correlated with the DTD such that listeners showing smaller performance decrement under dual attention had lower SNRs. Neither phoneme discrimination nor DTT_{VN} were correlated with any cognitive measure. There were also no correlations between any of the speech perception tests and age.

In addition to these results, Supplementary Tables S1 and S2 report the full set of (i) bivariate correlation coefficients, and (ii) all correlations with BEA partialled out. The partial correlations led to broadly similar results as seen with simple correlations. Noteworthy were three differences. First, ASL sentence perception was now negatively correlated with NVIQ

TABLE 1 | Mean, SE, and range for all measurements.

Domain	Function	Tests		Mean	SE	Range
Listening activity		PD (%) (n = 43)		65.35	1.43	54 to 99
		DTT_{VS} (SNR)		−6.60	0.23	−8.50 to −1.20
		DTT_{VN} (SNR)		−6.99	0.25	−9.20 to −1.00
		ASL (SNR)		1.34	0.60	−5.00 to 11.25
Hearing		BEA (0.5–4 kHz; dB HL)		32.50	0.90	21.25 to 45.00
Cognition	NVIQ	Matrix Reasoning		22.14	0.79	7 to 30
	Working memory (WM)	Digit Span	Forward	9.68	0.31	7 to 14
			Backward	6.68	0.29	4 to 11
		VLM (n = 40)	Accuracy (Hits)	5.65	0.41	0 to 10
			Speed (RTms)	644.30	22.10	0 to 881
	Attention	TEA (n = 43)	Subtest 6	3.40	0.11	2.10 to 4.80
			Subtest 7	4.69	0.18	2.60 to 7.0
			DTD	1.28	0.15	−0.10 to 3.70
Self-report	Health-related quality of life (HRQoL)	EQ-5D (n = 42)	General	1.26	0.05	1 to 2
			Hearing-specific	1.31	0.05	1 to 2
	Self-report of hearing	ALDQ		2.10	0.04	1.52 to 2.61
		GHABP (%)		32.81	2.76	4.17 to 62.50
		SSQ		6.29	0.22	3.07 to 9.49

PD, Phoneme discrimination; DTT, Digit Triplet Test with variable speech (DTT_{VS}) or variable noise (DTT_{VN}); ASL, Adaptive Sentence List; BEA, better ear average$_{(0.5-4\ kHz)}$; NVIQ, non-verbal intelligence quotient; VLM, visual letter monitoring; RT, reaction time; TEA, Test of Everyday Attention; DTD, dual-task decrement; HRQoL, Health related quality of life; ALDQ, Auditory Lifestyle and Demand Questionnaire; GHABP, Glasgow Hearing Aid Benefit Profile; SSQ, Speech, Spatial and Qualities of Hearing. When deviant from n = 44, n is noted for the particular test.

with a higher NVIQ score indicating better intelligibility and thus lower achieved SNR. Second, the previously significant negative correlation with Digit Span backward was now marginal. Third, the previously marginal positive correlation between DTT_{VS} and the DTD became significant. In summary, ASL and DTT_{VS} were associated with various tests of cognition, with a largely similar correlational pattern for bivariate and partial correlations.

In summary, in concordance with the prediction, the results show correlations between speech perception and cognitive tests, particularly in the cases of sentence perception (ASL) and DTT_{VS}. Although speech perception was also correlated with hearing sensitivity, the fundamental pattern of correlation between cognition and speech did not change much when hearing loss was partialled out. This suggests a genuine role of cognition for speech perception performance.

It is also interesting to note that the significant difference between correlation coefficients is often between ASL and DTT_{VS} for a particular cognitive variable. For instance in Supplementary Table S2, a significant correlation exists between ASL and both Matrix Reasoning and TEA6. The same is not true between DTT_{VS} and Matrix Reasoning and TEA6. In addition to being significant, the correlation coefficient between these cognitive variables and ASL was also significantly larger than that between the same cognitive variables and DTT_{VS}. Similarly, for TEA7, the correlation was significant with DTT_{VS} but not ASL, and the difference in correlation coefficient was in itself significant. Hence, while both DTT_{VS} and ASL correlate with cognitive measures, the correlation profile for these two speech perception tests differs, suggesting their cognitive requirements are different.

Prediction 1.2. The Contribution of Cognition will Increase as the Complexity of the Speech Perception Task Increases

Principal components analysis (PCA)

The principal component solutions based on the shared variance between all seven cognitive tests are shown in **Table 2**. Extracting a single principal component explained 40% of shared variance [KMO: 0.71, Bartlett: χ^2 (21) = 74.8, $p < 0.0001$] and showed substantial correlations with Matrix Reasoning, Digit Span forward and backward, VLM accuracy, and TEA 6 and 7 thereby representing a broad cognitive factor that includes non-verbal intelligence, WM, and attention. Only VLM Speed representing processing speed was not well represented by this latent factor.

Alternatively, aiming for the solution with the greatest amount of explained variance by extracting all factors with eigenvalue > 1 resulted in two factors and a total explained variance of 63% [KMO: 0.71, Bartlett: χ^2 (21) = 74.8, $p < 0.0001$)]. Factor 1, representing 33% of variance in cognitive performance, was most highly correlated with WM while Factor 2, explaining 30% of cognitive performance variance, loaded most highly on NVIQ and attention. Processing speed did not load highly on either factor. In the following, the single latent factor is referred to as General Cognition (Cogn) factor, and Factor 1 in the two-factor solution as WM factor, and Factor 2 in the two-factor solution as Attention (Att) factor.

Hierarchical regression analysis

Both the single Cogn factor and the two WM and Att factors were used as independent predictors in forward stepwise regression analyses on the four speech perception tests where they were always entered in a second step after hearing thresholds.

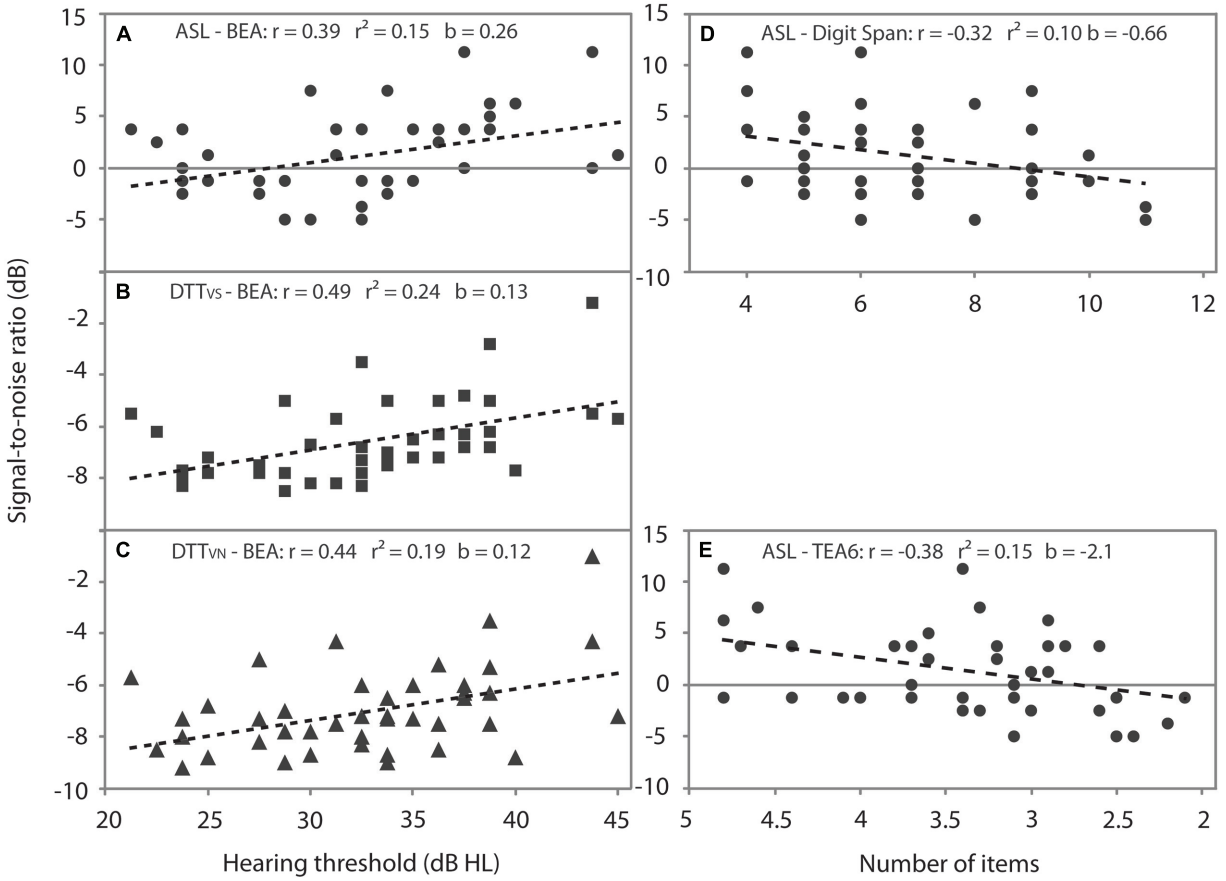


FIGURE 1 | Correlation coefficients (*r*), proportion of explained variance (r^2) and predictor values (*b*) for all significant correlations. Correlations with BEA are shown for **(A)** ASL, **(B)** DTT_{VS}, **(C)** DTT_{VN}, **(D)** correlations between ASL and Digit Span backward, and **(E)** ASL and TEA6. The dotted line shows the line of best fit. Acronyms as for **Table 1**.

TABLE 2 | Factor loadings for all cognitive tests for the two principal component analysis.

			Single factor solution	Two-factor solution	
			General Cognition (Cogn; 40%)	Working Memory (33%)	Attention (Att; 30%)
NVIQ	Matrix Reasoning		0.74	0.41	0.67
WM	Digit Span	Forward	0.76	0.83	0.18
		Backward	0.75	0.85	0.14
	VLM	Accuracy	0.60	0.80	−0.04
		Speed (RT ms)	−0.04	0.28	−0.41
Attention	TEA	Subtest 6	0.59	0.09	0.83
		Subtest 7	0.61	0.13	0.82

*Explained variance by each factor is in brackets. Acronyms as for **Table 1**. The loadings of the cognitive tests contributing most to a particular factor are shaded.*

The results of these analyses are reported in **Table 3**. For Phoneme discrimination, neither hearing nor cognition, either as single factor or two factors contributed significantly to the performance. For the two Digit Triplet tests, only hearing made a highly significant contribution, while cognition, whether entered as one (Cogn) or two (WM, Att) latent factors, did not. For Sentence perception, both hearing and cognition made significant contributions. Intriguingly, when the two latent cognitive factors WM and Att were entered separately into the model (M2), only Att made a significant contribution to Sentence perception suggesting that it was the attentional component in the cognitive tasks that drove the link with performance for this speech perception test.

This result extends the correlational results and suggests different predictive patterns of hearing and cognition for the speech perception tests. Specifically, it shows that the role of

TABLE 3 | Results for two forward stepwise regression models carried out for each speech perception test.

			R	R^2	adj *R*	SE	R^2 change	*F* change	*df*1	*df*2	Significance *f* change
PD	M1&M2	BEA	0.29	0.08	0.06	9.51	0.08	3.22	1	36	0.08
	M1	Cogn	0.32	0.10	0.05	9.55	0.02	0.72	1	35	0.40
	M2	WM	0.33	0.11	0.06	9.52	0.02	0.96	1	35	0.33
		Att	0.33	0.11	0.03	9.66	0.001	0.02	1	34	0.90
DTT_{VS}	M1&M2	BEA	0.49	0.24	0.22	1.43	0.24	11.76	1	37	0.002
	M1	Cogn	0.51	0.26	0.21	1.44	0.01	0.66	1	36	0.42
	M2	WM	0.50	0.25	0.21	1.45	0.01	0.25	1	36	0.62
		Att	0.51	0.26	0.19	1.46	0.01	0.44	1	35	0.51
DTT_{VN}	M1&M2	BEA	0.47	0.22	0.20	1.53	0.22	10.7	1	37	0.002
	M1	Cogn	0.48	0.23	0.19	1.54	0.006	0.28	1	36	0.60
	M2	WM	0.48	0.23	0.18	1.55	0.002	0.08	1	36	0.78
		Att	0.48	0.23	0.17	1.56	0.005	0.24	1	35	0.63
ASL	M1&M2	BEA	0.39	0.15	0.13	3.84	0.15	6.64	1	37	0.01
	M1	Cogn	0.51	0.27	0.22	3.63	0.11	5.51	1	36	0.02
	M2	WM	0.42	0.18	0.13	3.84	0.03	1.12	1	36	0.30
		Att	0.54	0.29	0.23	3.61	0.12	5.68	1	35	0.02

In all models hearing was entered first and cognition second, therefore results for hearing were identical regardless of how cognition was entered and is only reported once (M1&M2). In Model 1 (M1) for each speech perception test cognitive performance was entered as a single factor (Cogn). In Model 2 (M2) cognitive performance was entered as two separate factors representing WM and Attention. Acronyms as for ***Table 1****. Significant results ($p < 0.05$) are shaded.*

cognition was only predictive for performance differences in sentence perception. The main limitation of this approach is that the four speech perception tests are examined in separate statistical models. CCA examines whether there are correlations between two sets of variables and checks how many dimensions are shared between them. In this case hearing and cognition comprised one set, and the four speech perception tests the other set.

Canonical correlational analyses

The two sets that were compared comprised hearing, represented by BEA, and cognition, represented by the single PCA factor solution (Cogn), in Set 1 and the four speech perception tests in Set 2. The overall multivariate model, based on 38 cases, indicates that there is evidence for an overall relationship between the two sets of variables (Wilks' lambda, $p = 0.05$). Univariate regression analyses within the CCA model replicate the earlier hierarchical regression analyses by showing that performance on the DTT_{VS} [$F(2,35) = 6.04$, $p = 0.006$], DTT_{VN} [$F(2,35) = 5.12$, $p = 0.01$], and ASL [$F(2,35) = 6.12$, $p = 0.005$], but not on Phoneme discrimination [$F(2,35) = 1.95$, $p = 0.16$], showed significant contributions of at least one of the two predictor variables hearing and cognition. For the two digit triplet tests these contributions were due to hearing only ($p = 0.03$), whereas for sentence perception, both hearing ($p = 0.027$) and cognition ($p = 0.027$) contributed. The first canonical root explained 31% of shared variance, the second 9%, however, only the first root was significant (both canonical roots included: $F(8,64) = 2.11$, $p = 0.05$; first canonical root removed: $F(3,33) = 1.10$, $p = 0.36$). The correlations and canonical coefficients (loadings) for both solutions are included in Supplementary Table S3. Examination of the loadings suggests that hearing contributes about twice as much to the first root as cognition, and that the contribution of hearing and cognition were in opposite directions for the second root. Sentence perception was more affected by both root solutions than the other three speech perception tests.

In summary, based on all the statistical testing, a converging picture emerges in which cognitive tests differ in the extent to which they correlate with speech perception tests that vary in complexity. When cognition together with hearing, is considered as a predictor for speech perception performance, it only has a significant effect for sentence perception. This is true whether it is modeled as a unified variable or as a variable with subcomponents for WM and attention. Moreover, it is the attentional component of cognition that is crucial. Lastly, while the direct comparison of hearing and cognition for all four speech perception tests was limited by the small number of cases, and thus any results can only indicate tendencies, the CCA showed that the best root solution comprised both contributions from hearing and cognition and that this root was most important for modeling performance on the sentence perception test (ASL).

Prediction 1.3. Where Procedural Differences in Identifying SNR Occur while the Speech and Background Signals are Identical, We Expect Comparable Associations with Cognition if these Associations are Driven by Signal Complexities and not Procedural Differences

Supplementary Tables S1–S3 and **Table 3** suggest very similar results for DTT_{VS} and DTT_{VN} in relation to cognition. In Supplementary Tables S1 and S2, the correlation coefficients between DDT_{VS} or DTT_{VN} and a particular cognitive test are always almost identical. For correlation differences of this size to reach significance, at least 250 but often several 1000 participants would need to be tested. Similarly, in the CCA the weighting of the root factor, that is the effect of hearing and cognition, is very similar for the two types of digit triplet test (0.20 and 0.32). Lastly, in the stepwise regression analyses reported in **Table 3** both types

of digit triplet test showed the same predictive pattern for hearing (yes) and cognition (no). Hence, we conclude that there were no distinguishing features in these analyses to suggest that the relationship with cognition differs between DTT_{VS} and DTT_{VN}.

Aim 2: Assessing the Relationship between Speech Perception Performance and Self-Reported Outcomes

Prediction 2.1. Hearing-Specific Questionnaires will Demonstrate a Greater Association with Speech Perception Performance than Generic Health Measures

Correlational analyses

Simple Pearson product-moment correlations for the association between self-report measures and the four speech perception tests are shown in **Table 4**. The results show that the general HRQoL questions (general EQ5-5D) were not correlated with performance on any of the speech perception tests. In contrast, hearing and communication-specific measures (hearing-specific EQ-5D, ALDQ, GHABP, and SSQ) were significantly associated with some, but not all, speech perception tests. Hence, only questionnaires that assessed hearing-related aspects of self-report correlated with behavioral measures of speech perception.

Prediction 2.2. Correlations with Speech Perception Performance will be Largest for Questionnaires that Capture Aspects of Listening Important for that Particular Speech Perception Test

Correlational analyses – differences between tests

Table 4 also shows that DTT_{VN} had the greatest number of significant correlations with self-report questionnaires, in particular with the hearing-specific EQ-5D and the hearing-specific questionnaires (ALDQ, GHABP, and SSQ). In contrast, Phoneme discrimination was not correlated with any self-report questionnaires. Both DTT_{VS} and Sentence perception were only each correlated with one self-report scale (SSQ and hearing-specific EQ-5D, respectively). A direct comparison of correlation sizes between speech perception and self-report measures ('Diff significant') showed that even though DTT_{VN} had numerous significant correlations with self-report measures, the coefficients were not significantly greater than those for the ASL or DTT_{VS}, except for ASL in the case of GHABP. Hence, it is not clear whether one particular SiN test captures self-report significantly better than other speech perception tests.

Canonical correlational analyses

The four speech perception tests were entered as one set of variables, while the hearing-specific EQ-5D, the ALDQ, GHABP, and SSQ were entered as the other set. The overall multivariate model, based on 41 cases, indicated that there was evidence for an overall relationship between the two sets of variables (Wilks' lambda, $p = 0.005$). Univariate regression analyses within the CCA model indicated that only performance on Phoneme discrimination was not significantly related to self-report, while performance on all other speech perception tests was significantly related to self-report (DTT_{VS}: $p = 0.016$; DTT_{VN}: $p = 0.005$; ASL: $p = 0.025$). The first canonical root explained 38% of shared variance, the second 26%, the third 10%, and the fourth 9%, with only the first two roots being significant [all canonical correlations included: $F(16,101) = 2.37$, $p = 0.005$; first root removed: $F(9,83) = 2.09$, $p = 0.04$]. The correlations and canonical coefficients for the significant root solutions 1 and 2 are shown in Supplementary Table S4. Examination of the loadings suggests a picture similar to that presented by the correlations reported in **Table 4**. The first canonical root suggests that lower scores on hearing-specific EQ-5D and higher (i.e., richer) scores on self-rated sound environments are related to higher SNR in the DTT_{VN}. This replicates the negative correlation between hearing-specific EQ-5D and DTT_{VN}, and the positive correlation between ALDQ and DTT_{VN}. The second canonical root suggests that better self-rated activity and participation scores are related to lower SNRs in the DTT_{VN}. This replicates the negative correlation between GHABP and DTT_{VN}.

DTT_{VN} showed the richest pattern of correlations with self-report questionnaires, although this difference in pattern was to some extent difficult to establish in terms of significant

TABLE 4 | Pearson product-moment correlations between each of four speech perception tests and self-report questionnaires.

			Correlation				Diff. significant		
			PD	**DTT_{VS}**	**DTT_{VN}**	**ASL**	**DTT_{VS}**	**DTT_{VN}**	**ASL**
Health-related quality of life (HRQoL)	EQ-5D	General	0.10	0.11	−0.08	−0.17	DTT_{VN}[a] ASL[b]	DTT_{VS}[a]	DTT_{VS}[b]
		N	41	42	42	42			
		Hearing-specific	0.10	−0.26	−0.38*	−0.42**	PD[b]	PD[b]	PD[b]
		N	41	42	42	42			
Self-report of hearing	ALDQ		0.25	0.25	0.35*	0.26			
		N	43	44	44	44			
	GHABP		0.04	−0.26	−0.33*	−0.02	ASL[a]	PD[b], ASL[b]	DTT_{VS}[a] DTT_{VN}[b]
		N	43	44	44	44			
	SSQ		−0.06	−0.37*	−0.29*	−0.25	PD[b]		
		N	43	44	44	44			

*Acronyms as for **Table 1**. Significant two-tailed correlations are shaded. *$p < 0.05$, **$p < 0.01$, [a]$p_{(one-sided)} < 0.05$, [b]$p_{(two-sided)} < 0.05$.*

differences in correlation size. This difference in association between speech perception tests and questionnaires was also reflected in the canonical correlations. Despite differences being small, the overall pattern of results nevertheless suggests that speech perception tests differ in how closely their performance is associated with aspects of self-reported hearing, and that performance on the DTT_{VN} showed the closest correspondence with all the hearing-related self-report scales.

Prediction 2.3. Procedural Differences in Identifying SNR for Speech Perception Performance may Lead to Different Associations with Self-report Scales. In Particular, Increasing the Level of Background Noise to Reduce Perceptual Accuracy may be Uniquely Associated with Functioning in Challenging Auditory Environments

This hypothesis is assessed by comparing the differences in correlation between self-report scales and DTT_{VS} or DTT_{VN}, respectively. **Table 4** shows that the differences in correlation between self-report scales and the two speech perception tests are small. For the hearing-specific EQ-5D, ALDQ, and GHABP the differences are 0.12, 0.10, and 0.08 which equates to a small effect. In the context of this study more than 80 participants would be required for an effect of this magnitude to reach significance. Nevertheless, the canonical correlations suggest the involvement of particularly DTT_{VN} in several correlations of different aspects of the speech perception.

Discussion

Listening can be assessed behaviorally with speech perception tests or subjectively with self-report measures. Which measure is chosen to assess an outcome, either in clinical or research evaluations, depends on many factors including availability, familiarity, and popularity of a particular measure. Less consideration might be given to either the specific aspect of listening that is assessed by a particular test or questionnaire, or the contribution of cognitive functioning to speech perception performance. This investigation considered these relationships to help inform outcome selection for clinical and research purposes.

We assessed the relationship between measures of speech perception and hearing, cognition, and self-reported outcomes. Speech perception tests varied in complexity from low (phonemes in quiet) to medium (words in steady-state speech-shaped noise) to high (sentences in 8 Hz modulated noise). Cognitive tests either emphasized the storage and processing of information (WM), or attention and cognitive control. Information storage and processing capacities were measured with digit span tasks (forward and backward) and a VLM task, while attention and cognitive control was measured by means of focused and divided attention tasks. We also assessed the effect of the protocol for changing the SNR in one of the speech tasks (Digit Triplet Test) by varying either the speech or the noise. This allowed us to assess whether the procedure affected either the extent of cognitive contributions to the speech task, or the extent to which speech perception performance correlated with self-reported aspects of hearing. In the following, each hypothesis and associated results is considered in turn.

Assessing the Relationship between Speech Perception Performance and Cognition

Prediction 1.1. Speech Perception Performance will be Associated with Cognition, and this will be Moderated by Hearing Sensitivity

Initial correlation analyses showed some correlations between speech perception and cognitive performance. This pattern remained largely unchanged even when hearing loss was taken into account, despite the fact that hearing loss had a significant influence on the speech perception results. Age did not independently contribute to the speech perception results, possibly because the age range of the participants was restricted (50–74 years).

The influence of hearing loss on speech perception is well documented in the literature (e.g., Humes and Roberts, 1990; van Rooij and Plomp, 1990; Humes and Dubno, 2010) and the results of this study fit within this body of evidence. That cognition also presented as a considerable factor for speech perception performance in some tests, above and beyond hearing loss, is also in accordance with previous results (e.g., Akeroyd, 2008; Houtgast and Festen, 2008; Humes et al., 2013). Finally, studies have also previously shown that the contribution of hearing and cognition to speech perception performance varies depending on the background in which the speech task is presented, with adverse noise conditions more likely to invoke cognitive processes than listening in quiet (e.g., van Rooij and Plomp, 1990; Wingfield et al., 2005; Rönnberg et al., 2010). However, the complexity of a listening situation can vary in more ways than just the presence of absence of background noise. Thus, the second prediction was investigated to assess how the contribution of cognition changed depending on the listening situation.

Prediction 1.2. The Contribution of Cognition will Increase as the Complexity of the Speech Perception Task Increases

The complexity of the listening situation in the current study is determined by (i) the target speech, which comprised phonemes, words or sentences, (ii) the background, which was steady-state and 8-Hz modulated noise, and (iii) the listening task itself, which included recognition and comprehension. How these different aspects of the listening situation affect the relationship between cognitive processing and speech perception have so far inspired surprisingly little systematic research, apart from the general demonstration that correlations with cognitive processes are greater when listening to speech in adverse noise conditions than when listening in quiet (e.g., van Rooij and Plomp, 1990; Wingfield et al., 2005; Rönnberg et al., 2010). This study took a first step toward understanding if and how the contribution of cognitive components differed for various SiN conditions, and whether this depended on the exact pairing of cognitive subcomponent and complexity of listening situation.

The choice of cognitive subcomponents to be assessed was informed by previous work that had clearly demonstrated a role of WM for SiN perception (see Akeroyd, 2008 for a review). However, WM tests differ in respect to the emphasis they give to different subcomponents of cognition (storage, processing, inhibition, cognitive control) depending on the model they are based on. The current study tested all of these subcomponents. On a general note, the study showed that correlations between cognitive components and speech perception occurred mainly for the most complex speech perception test (sentences in 8 Hz modulated noise), while digit perception in steady-state noise showed only few correlations, and phoneme discrimination in quiet showed none. This result was also borne out in the hierarchical regression analyses where only performance on the sentence perception task was reliably associated with cognition.

Distinct cognitive profiles for different speech perception tests emerged, in particular for the sentence perception and DTT_{VS}. Supplementary Table S2 shows that not only performance on the NVIQ and focused attention tasks correlated significantly with sentence perception but also that this correlation was significantly higher than those for the same tests with DTT_{VS}. For the divided attention decrement, the situation was reversed in that this test only showed a significant correlation with $DTT_{VS,}$ which was significantly higher than with sentence perception. At this point we can only speculate why this might have happened as we did not systematically manipulate aspects of the listening task to assess whether it was the change in target speech (from digits to sentences) or the change in background noise (from steady-state to modulated noise) that led to this change in correlation profile. It may be the correlation between sentence perception and digit span occurred because the successful repetition of a sentence involved significant WM storage. It is also possible that focused attention on the words within a sentence was particularly beneficial because perception of words may result in successful inference of the rest of the sentence, whereas such an inference would not be possible for strings of single digits. Conversely, for digit triplet in noise, maybe successful listening meant being able to tolerate both signals, the digits and noise, rather than trying (and failing) to ignore the noise, and listeners who were best able to do this also had the smallest divided task decrement.

These data offer some initial suggestions that may help to reconcile the inconsistencies existing in the literature on the relationship between cognition and speech perception, and may thereby help to increase our understanding of the exact relationship between speech perception and cognition. The results suggest that the relationship between speech and cognition can be specific to the tests used, and thus simply referring to speech perception and cognition may ignore important distinctions. Being more specific about cognition and speech may help us understand why the reading span task, as a complex WM measure, correlates with speech perception when measured with sentences in noise (Desjardins and Doherty, 2013; Moradi et al., 2014) but not when measured with syllables (Kempe et al., 2012). Similarly, performance on the VLM task may predict performance on a particular word perception task (Gatehouse et al., 2003) but not on a sentence perception task (Rudner et al., 2008).

Lastly, when assessing the effect of WM and attention for cognition (Att) separately by means of latent principal component factors, it was the attention and NVIQ, rather than WM that were associated most closely with sentence perception performance. This result contrasts with previous studies which have shown a clear correlation between WM and SiN perception in older listeners (Humes et al., 2006; Rudner et al., 2008).

Prediction 1.3. Where Procedural Differences in Identifying SNR Occur while the Speech and Background Signals are Identical, we Expect Comparable Associations with Cognition if these Associations are Driven by Signal Complexities and not Procedural Differences

An interesting dichotomy of results emerged: ASL and DTT_{VS} which both changed SNR in the same way (constant noise level and adjusted speech) but used different speech material (sentences and words) showed statistically reliable differences in their cognitive profiles (i.e., their correlations with specific cognitive tests). Conversely, DTT_{VS} and DTT_{VN}, which both used different methods to adjust SNR, but also used the same speech material and background sounds, showed similar cognitive profiles. It might be argued that the similarity in results between DTT_{VS} and DTT_{VN} was due to insufficient power rather than the true absence of an effect. However, the significant differences between ASL and DTT_{VS} showed that the effects in the data were strong enough to show significant differences when they existed. Moreover, power analyses based on the current effect sizes showed that for most profile differences several 100 data points would have been needed to show significant differences. Therefore we conclude that our results were consistent with the prediction, and that both methods of setting SNRs place similar cognitive demands on the listener and are equally suited for setting SNR if cognitive demand is the main concern.

Assessing the Relationship between Speech Perception Performance and Self-Reported Outcomes

Prediction 2.1. Hearing-Specific Questionnaires will Demonstrate a Greater Association with Speech Perception Performance than Generic Health Measures

Questionnaires that assess activity and participation relating to hearing and communication correlated more highly with speech perception outcomes than general HRQoL questionnaires. These results are consistent with other studies (Joore et al., 2002; Stark and Hickson, 2004; Chisolm et al., 2007) and this prediction.

Prediction 2.2. Correlations with Speech Perception Performance will be Largest for Questionnaires that Capture Aspects of Listening Important for that Particular Speech Perception Test

Similar to the cognitive results, different patterns of correlation also existed between self-report measures and speech perception tests. Phoneme discrimination correlated least with self-report

measures. At this point we cannot say whether this result occurred because of the low complexity of the speech material or the lack of background noise, or indeed both. All other speech perception tests showed correlations with at least one self-report outcome. Although DTT_{VN} showed the richest pattern of significant correlations with self-report measures, the differences in correlation to the other speech perception tests involving at least words or sentences only became significant for one questionnaire (i.e., GHABP), and only in contrast to one speech perception test (i.e., ASL). In summary, these results would suggest that these speech perception tests all measure similar aspects of self-reported experiences but that these aspects are represented most strongly in the DTT_{VN}.

Prediction 2.3. Procedural Differences in Identifying SNR for Speech Perception Performance may Lead to Different Associations with Self-Report Scales. In Particular, Increasing the Level of Background Noise to Reduce Perceptual Accuracy may be Uniquely Associated with Functioning in Challenging Auditory Environments

One particularly interesting aspect of the study was the administration of the same speech task, the DTT, with two different administration protocols and the resulting changes in the correlation with self-reported outcomes. The results showed that administering the task with variable noise (DTT_{VN}) was significantly associated with aspects of communication (hearing-specific EQ-5D), ALDQ, communication (GHABP), and SSQ. However, administering the task with variable speech (DTT_{VS}) was only significantly associated with the SSQ. Moreover, in the CCA, DTT_{VN} contributed substantially more to the first and second canonical root than DTT_{VS}, suggesting that DTT_{VN} is more likely to play a prominent role in hearing and communication functions. This is relevant to the way in which the DTT was administered, and highlights the fact that practitioners and researchers alike should think about their question of interest before deciding for a particular test. If aspects of speech perception are of most interest then fixing the noise level and varying the speech appears most effective. However, if aspects of communication and participation restriction of the listening experience are of interest, then choosing to keep the level of the speech constant and varying the noise might be more appropriate. These results are also interesting in the light of previous research, where some studies have used variable speech (Plomp and Mimpen, 1979; Smits et al., 2004; George et al., 2006; Jansen et al., 2010; Vlaming et al., 2011), while others have used variable noise (Mayo et al., 1997; Rogers et al., 2006), with one study even using both methods in the same experiment (Smits et al., 2013). If communication ability and noise tolerance beyond intelligibility is a consideration then researchers need to choose deliberately between the two SNR methods.

Limitations

There are a number of limitations to this investigation. First, this study was designed as an auditory training intervention trial. Therefore the measures were included for the purpose of assessing the intervention, and not specifically selected for the purposes of the current evaluation. As such, speech and cognitive outcomes were limited to the outcomes of that study, and were not chosen specifically to represent a fully factorial combination of the complexities of target speech and background noise. Instead they were meant to sample broadly across the continuum of listening situations with varying complexities in foreground and background simultaneously. As a result, changes in correlations between cognitive function and speech perception cannot be unambiguously attributed to changes in the complexity of the target speech. Future purpose-designed studies will enable a finer-grained analysis of the issues raised in this investigation and investigate in greater detail the complexity of the foreground and background signal to listening demands.

Another consequence of the intervention trial design is the fact that the number of participants ($n = 44$), while large for a training study, is rather small for the type of analyses performed here. This limits the power and generalizability of the results. The coarse differentiation of speech perception test complexity and the relatively small number of participants makes this study strictly exploratory.

Third, the inherent nature of a speech perception test dictates that the speech content is unlikely to be highly relevant to the individual, nor particularly interesting. This may therefore impact on an individual's motivations to pay attention and actively listen to the speech content (see Henshaw et al., in press for an overview).

Fourth and finally, the participants in this study were adults with mild SNHL who did not wear hearing aids. Thus, this investigation adds to research on the relationship between cognition and self-report measures to different speech perception tests in un-aided listening (Cox and Alexander, 1992; Humes et al., 2013). This stand-alone examination cannot tell us how these relationships may change once hearing intervention occurs, e.g., once hearing aids are fitted.

Conclusion

The results of this study show that different speech perception tests engage cognition to different extents, and reflect different subjective aspects of the self-reported listening experience. These results suggest that practitioners and researchers should think carefully about the objective outcome measures they choose as different speech and cognitive tests will highlight different aspects of listening and engage different cognitive processes. One way in which this could be useful for audiological practice is to choose a speech perception test that highlights those aspects of communication and participation that the patient indicated as being important and/or difficult for them. Alternatively, tests could be specifically chosen to maximize or minimize cognitive influences, which might put a listener at an advantage or a disadvantage. Finally, to assess change in speech perception performance as a result of an intervention, researchers or clinicians should select speech perception tests that are associated

with the intended mechanism of benefit of that intervention in order to adequately detect any associated change in performance (see Ferguson and Henshaw, 2015).

Author Contributions

MF designed the study. AH analyzed and interpreted the data. AH wrote the manuscript. AH, MF, and HH contributed to critical discussions. AH and MF revised the manuscript. All authors approved the final version of the manuscript for publication. All authors agree to be accountable for all aspects of the work and in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgments

The authors would like to thank Oliver Zobay for support with statistical analyses. This paper presents independent research funded by the National Institute for Health Research (NIHR) Biomedical Research Unit Programme. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00782/abstract

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

Arlinger, S., Gatehouse, S., Kiessling, J., Naylor, G., Verschuure, H., and Wouters, J. (2008). The design of a project to assess bilateral versus unilateral hearing aid fitting. *Trends Amplif.* 12, 137–144. doi: 10.1177/1084713808316171

Barrouillet, P., Bernardin, S., and Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *J. Exp. Psychol. Gen.* 133, 83–100. doi: 10.1037/0096-3445.133.1.83

Boothroyd, A. (2007). Adult aural rehabilitation: what is it and does it work? *Trends Amplif.* 11, 63–71. doi: 10.1177/1084713807301073

British Society of Audiology. (2011). *Recommended Procedure for the Pure Tone Air and Bone Conduction Threshold Audiometry with and without the Use of Masking and Determination of Uncomfortable Loudness Levels [Online]*. Available at: http://www.thebsa.org.uk/wp-content/uploads/2014/04/BSA_RP_PTA_FINAL_24Sept11_MinorAmend06Feb12.pdf (accessed February 11, 2015).

Chisolm, T. H., Johnson, C. E., Danhauer, J. L., Portz, L. J. P., Abrams, H. B., Lesner, S., et al. (2007). A systematic review of health-related quality of life and hearing aids: final report of the American academy of audiology task force on the health-related quality of life benefits of amplification in adults. *J. Am. Acad. Audiol.* 18, 151–184. doi: 10.3766/jaaa.18.2.7

Cox, R. M., and Alexander, G. C. (1992). Maturation of hearing aid benefit: objective and subjective measurements. *Ear Hear.* 13, 131–141. doi: 10.1097/00003446-199206000-00001

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Danermark, B., Cieza, A., Gangé, J. P., Gimigliano, F., Granberg, S., Hickson, L., et al. (2010). International classification of functioning, disability, and health core sets for hearing loss: a discussion paper and invitation. *Int. J. Audiol.* 49, 256–262. doi: 10.3109/14992020903410110

Desjardins, J. L., and Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear Hear.* 34, 261–272. doi: 10.1097/AUD.0b013e31826d0ba4

Engle, R. W., and Kane, M. J. (2003). "Executive attention, working memory capacity, and a two-factor theory of cognitive control," in *Psychology of Learning and Motivation*, ed. B. Ross (San Diego, CA: Elsevier), 145–199.

Ferguson, M. A., and Henshaw, H. (2015). Auditory training can improve working memory, attention and communication in adverse conditions for adults with hearing loss. *Front. Psychol.* 6:556. doi: 10.3389/fpsyg.2015.00556

Ferguson, M. A., Henshaw, H., Clark, D., and Moore, D. (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50–74 year olds with mild hearing loss. *Ear Hear.* 35, e110–e121. doi: 10.1097/AUD.0000000000000020

Gatehouse, S. (1999). Glasgow hearing aid benefit profile: derivation and validation of a client-centered outcome measure for hearing aid services. *J. Am. Acad. Audiol.* 10, 80–103.

Gatehouse, S. (2003). Rehabilitation: identification of needs, priorities and expectations, and the evaluation of benefit. *Int. J. Audiol.* 42, 77–83. doi: 10.3109/14992020309074648

Gatehouse, S., Elberling, C., and Naylor, G. (1999). "Aspects of auditory ecology and psychoacoustic function as determinants of benefits from and candidature for non-linear processing hearing aids," in *Auditory Models and Non-Linear Hearing Instruments*, eds A. N. Rasmussen, P. A. Osterhammel, T. Andersen, and T. Poulsen (Denmark: The Danavox Jubilee Foundation), 221–233.

Gatehouse, S., Naylor, G., and Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Int. J. Audiol.* 42, S77–S85. doi: 10.3109/14992020309074627

Gatehouse, S., and Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *Int. J. Audiol.* 43, 85–99. doi: 10.1080/14992020400050014

George, E. L. J., Festen, J. M., and Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 120, 2295–2311. doi: 10.1121/1.2266530

Granberg, S., Dahlström, J., Möller, C., Kähäri, K., and Danermark, B. (2014). The ICF Core Sets for hearing loss–researcher perspective. Part I: systematic review of outcome measures identified in audiological research. *Int. J. Audiol.* 53, 65–76. doi: 10.3109/14992027.2013.851799

Grutters, J. P. C., Joore, M. A., Van Der Horst, F., Verschuure, H., Dreschler, W. A., and Anteunis, L. J. C. (2007). Choosing between measures: comparison of EQ-5D, HUI2 and HUI3 in persons with hearing complaints. *Qual. Life Res.* 16, 1439–1449. doi: 10.1007/s11136-007-9237-x

Henshaw, H., Mccormack, A., and Ferguson, M. A. (in press). Intrinsic and extrinsic motivation drives computer-based auditory training uptake, engagement, and adherence for people with hearing(loss). *Front. Aud. Cogn. Neurosci.*

Hickson, L., and Scarinci, N. (2007). Older adults with acquired hearing impairment: applying the ICF in rehabilitation. *Semin. Speech Lang.* 28, 283–290. doi: 10.1055/s-2007-986525

Houtgast, T., and Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *Int. J. Audiol.* 47, 287–295. doi: 10.1080/14992020802127109

Humes, L. E., and Dubno, J. R. (2010). "Factors affecting speech understanding in older adults," in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (New York, NY: Springer), 211–258.

Humes, L. E., Garner, C. B., Wilson, D. L., and Barlow, N. N. (2001). Hearing-aid outcome measures following one month of hearing aid use by the elderly. *J. Speech Lang. Hear. Res.* 44, 469–486. doi: 10.1044/1092-4388(2001/037)

Humes, L. E., Kidd, G. R., and Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding

among older adults. *Front. Syst. Neurosci.* 7:55. doi: 10.3389/fnsys.2013.00055

Humes, L. E., Lee, J. H., and Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *J. Acoust. Soc. Am.* 120, 2926–2937. doi: 10.1121/1.2354070

Humes, L. E., and Roberts, L. (1990). Speech-recognition difficulties of the hearing-impaired elderly: the contributions of audibility. *J. Speech Hear. Res.* 33, 726–735. doi: 10.1044/jshr.3304.726

Jansen, S., Luts, H., Wagener, K. C., Frachet, B., and Wouters, J. (2010). The French digit triplet test: a hearing screening tool for speech intelligibility in noise. *Int. J. Audiol.* 49, 378–387. doi: 10.3109/14992020903431272

Joore, M. A., Potjewijd, J., Timmerman, A. A., and Auteunis, L. J. (2002). Response shift in the measurement of quality of life in hearing impaired adults after hearing aid fitting. *Qual. Life Res.* 11, 299–307. doi: 10.1023/A:1015598807510

Kempe, V., Thoresen, J. C., Kirk, N. W., Schaeffler, F., and Brooks, P. J. (2012). Individual differences in the discrimination of novel speech sounds: effects of sex, temporal processing, musical and cognitive abilities. *PLoS ONE* 7:e48623. doi: 10.1371/journal.pone.0048623

Koelewijn, T., Zekveld, A. A., Festen, J. M., Ronnberg, J., and Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *Int. J. Otolaryngol.* 2012:865731. doi: 10.1155/2012/865731

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375

MacLeod, A., and Summerfield, Q. (1990). A procedure for measuring auditory and audio-visual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. *Br. J. Audiol.* 24, 29–43. doi: 10.3109/03005369009077840

Mattys, S., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006

Mayo, L. H., Florentine, M., and Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *J. Speech Lang. Hear. Res.* 40, 686–693. doi: 10.1044/jslhr.4003.686

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognit. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moradi, S., Lidestam, B., Saremi, A., and Ronnberg, J. (2014). Gated auditory speech perception: effects of listening conditions and cognitive capacity. *Front. Psychol.* 5:531. doi: 10.3389/fpsyg.2014.00531

Nabelek, A. K., Tucker, F. M., and Letowski, T. R. (1991). Toleration of background noises: relationship with patterns of hearing aid use by elderly persons. *J. Speech Hear. Res.* 34, 679–685. doi: 10.1044/jshr.3403.679

Plomp, R., and Mimpen, A. M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18, 43–52. doi: 10.3109/00206097909072618

Robertson, I. H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (1994). *The Test of Everyday Attention*. Bury St. Edmunds: Thames Valley Test Company.

Rogers, C. L., Lister, J. J., and Febo, D. M. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Appl. Psycholinguist.* 27, 465–485. doi: 10.1017/S014271640606036X

Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12, 26326–26329. doi: 10.4103/1463-1741.70505

Rudner, M., Foo, C., Sundewall-Thorén, E., Lunner, T., and Rönnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *Int. J. Audiol.* 47(Suppl. 2), S91–S98. doi: 10.1080/14992020802304393

Saunders, G. H., Chisolm, T. H., and Abrams, H. B. (2005). Measuring hearing aid outcomes–not as easy as it seems. *J. Rehabil. Res. Dev.* 42, 157–168. doi: 10.1682/JRRD.2005.01.0001

Schneider, B. A., and Pichora-Fuller, M. K. (2000). "Implications of perceptual deterioration for cognitive aging research," in *The Handbook of Aging and Cognition*, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Erlbaum), 155–219.

Schneider, B. A., Pichora-Fuller, M. K., and Daneman, M. (2010). "Effects of senescent changes in audition and cognition on spoken language comprehension," in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (New York, NY: Springer), 167–210.

Smits, C., Goverts, T., and Festen, J. M. (2013). The digits-in-noise test: assessing auditory speech recognition abilities in noise. *J. Acoust. Soc. Am.* 133, 1693–1706. doi: 10.1121/1.4789933

Smits, C., and Houtgast, T. (2005). Results from the Dutch speech-in-noise screening test by telephone. *Ear Hear.* 26, 89–95. doi: 10.1097/00003446-200502000-00008

Smits, C., Kapteyn, T. S., and Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *Int. J. Audiol.* 43, 15–28. doi: 10.1080/14992020400050004

Stark, P., and Hickson, L. (2004). Outcomes of hearing aid fitting for (older) people with hearing impairment and their significant others. *Int. J. Audiol.* 43, 390–398. doi: 10.1080/14992020400050050

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychol. Bull.* 87, 245–251. doi: 10.1037/0033-2909.87.2.245

Stucki, G., and Grimby, G. (2004). Foreword: applying the ICF in medicine. *J. Rehabil. Med.* 44, 5–6. doi: 10.1080/16501960410022300

The EuroQol Group. (1990). EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 16, 199–208. doi: 10.1016/0168-8510(90)90421-9

van Rooij, J. C. G. M., and Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II. Multivariate analyses. *J. Acoust. Soc. Am.* 88, 2611–2624. doi: 10.1121/1.399981

Vlaming, M. S. M. G., Kollmeier, B., Dreschler, W. A., Martin, R., Wouters, J., Grover, B., et al. (2011). HearCom: hearing in the communication Society. *Acta Acust. United Acust.* 97, 175–192. doi: 10.3813/AAA.918397

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (*WAIS-3®*), 3rd Edn. San Antonio, TX: Harcourt Assessment.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence* (*WASI*). San Antonio, TX: Harcourt Assessment.

WHO. (2001). *International Classification of Functioning, Disability and Health* (*ICF*). Geneva: WHO.

Wilson, R. H., Carnell, C. S., and Cleghorn, A. L. (2007). The Words-in-Noise (WIN) test with multitalker babble and speech-spectrum noise maskers. *J. Am. Acad. Audiol.* 18, 522–529.

Wingfield, A., and Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *J. Am. Acad. Audiol.* 18, 548–558. doi: 10.3766/jaaa.18.7.3

Wingfield, A., Tun, P. A., and Mccoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Heinrich, Henshaw and Ferguson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 16 September 2015
doi: 10.3389/fpsyg.2015.01394

Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss

Sherri L. Smith[1,2]* *and M. Kathleen Pichora-Fuller*[3,4,5,6]

[1] Audiologic Rehabilitation Laboratory, Auditory Vestibular Research Enhancement Award Program, Veterans Affairs Medical Center, Mountain Home, TN, USA, [2] Department of Audiology and Speech-Language Pathology, East Tennessee State University, Johnson City, TN, USA, [3] Department of Psychology, University of Toronto, Mississauga, ON, Canada, [4] Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada, [5] Rotman Research Institute, Baycrest Hospital, Toronto, ON, Canada, [6] Linneaus Centre HEAD, Linköping University, Linköping, Sweden

OPEN ACCESS

Edited by:
Isabelle Peretz,
Université de Montréal, Canada

Reviewed by:
Fatima T. Husain,
University of Illinois
at Urbana-Champaign, USA
Simona Brambati,
Université de Montréal, Canada

***Correspondence:**
Sherri L. Smith,
Audiologic Rehabilitation Laboratory,
Auditory Vestibular Research
Enhancement Award Program,
Veterans Affairs Medical Center,
Audiology 126, Mountain Home,
TN 37684, USA
sherri.smith@va.gov

Specialty section:
This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Psychology

Received: 03 March 2015
Accepted: 01 September 2015
Published: 16 September 2015

Citation:
Smith SL and Pichora-Fuller MK
(2015) Associations between speech
understanding and auditory and visual
tests of verbal working memory:
effects of linguistic complexity, task,
age, and hearing loss.
Front. Psychol. 6:1394.
doi: 10.3389/fpsyg.2015.01394

Listeners with hearing loss commonly report having difficulty understanding speech, particularly in noisy environments. Their difficulties could be due to auditory and cognitive processing problems. Performance on speech-in-noise tests has been correlated with reading working memory span (RWMS), a measure often chosen to avoid the effects of hearing loss. If the goal is to assess the cognitive consequences of listeners' auditory processing abilities, however, then listening working memory span (LWMS) could be a more informative measure. Some studies have examined the effects of different degrees and types of masking on working memory, but less is known about the demands placed on working memory depending on the linguistic complexity of the target speech or the task used to measure speech understanding in listeners with hearing loss. Compared to RWMS, LWMS measures using different speech targets and maskers may provide a more ecologically valid approach. To examine the contributions of RWMS and LWMS to speech understanding, we administered two working memory measures (a traditional RWMS measure and a new LWMS measure), and a battery of tests varying in the linguistic complexity of the speech materials, the presence of babble masking, and the task. Participants were a group of younger listeners with normal hearing and two groups of older listeners with hearing loss ($n = 24$ per group). There was a significant group difference and a wider range in performance on LWMS than on RWMS. There was a significant correlation between both working memory measures only for the oldest listeners with hearing loss. Notably, there were only few significant correlations among the working memory and speech understanding measures. These findings suggest that working memory measures reflect individual differences that are distinct from those tapped by these measures of speech understanding.

Keywords: hearing loss, speech understanding, aging, reading working memory, listening working memory, speech-in-noise

Introduction

For over a half century, researchers and clinicians have recognized that speech understanding difficulties are common amongst older listeners, particularly when speech is presented in a noisy background or when listeners have age-related hearing loss (e.g., Bocca and Calearo, 1963; Frisina and Frisina, 1997; Gates and Mills, 2005; Mills et al., 2006; Humes and Dubno, 2010). It is well known that both sensory and cognitive processes are independently and interactively involved in speech understanding (e.g., Committee on Hearing, Bioacoustics, and Biomechanics [CHABA], 1988). Research examining the interactions between sensory and cognitive processes has resulted in the emerging field of cognitive hearing science, with much of the recent work in this field focusing on the role that working memory plays in speech understanding in listeners who may have various degrees and types of hearing loss (Arlinger et al., 2009). Working memory is thought to be important for speech understanding because listeners must decode the incoming speech signal while relating the information to stored knowledge and anticipating the speech that is forthcoming (e.g., Daneman and Carpenter, 1980, 1983; Pichora-Fuller et al., 1995; Daneman and Merikle, 1996; Wingfield and Stine-Morrow, 2000; Akeroyd, 2008). When the audibility of the speech signal is reduced due to hearing loss or noise, then more working memory resources may need to be allocated when listeners are trying to comprehend the impoverished incoming speech signal (see also Rabbitt, 1968, 1991; van Rooij and Plomp, 1990; Wingfield, 1996; Lunner, 2003; Humes and Floyd, 2005; Foo et al., 2007; Rudner et al., 2007, 2011; Akeroyd, 2008; Parbery-Clark et al., 2009; Besser et al., 2013). Converging evidence from studies associating working memory measures to speech recognition measures (e.g., measures of how accurately words are repeated by listeners) suggests that inter-individual differences in working memory span explain a small portion of the variance and that listeners with high working memory span have better speech recognition in adverse listening conditions relative to those with low working memory capacity (see Akeroyd, 2008; Besser et al., 2013; and Humes et al., 2013 for reviews). Some studies, however, have been more successful than others in associating working memory and speech-recognition measures, perhaps in part due to the variations in the working memory and speech measures used.

Some researchers have suggested that when examining the associations between working memory and speech understanding, working memory measures should be presented in the visual domain to avoid potential sensory encoding issues associated with the auditory presentation of materials, particularly for listeners with hearing loss (e.g., Souza, 2012). Others have suggested, however, that because working memory is both domain- and modality-specific, it may be more appropriate to measure working memory using test materials presented in conditions that approximate the functional situation of interest (e.g., Pichora-Fuller et al., 1995; Baldwin and Ash, 2011; Besser et al., 2013; for a review of the issue of modality-specificity in testing auditory processing see Cacace and McFarland, 2013). In other words, to understand better the interplay of working memory and speech understanding in everyday listening conditions, it may be better to test working memory using auditory verbal stimuli. Both auditory and visual working memory tests have been used in recent studies, but the reading span measure has been the most commonly used in studies examining the association between working memory with speech recognition (see Akeroyd, 2008; Besser et al., 2013). Of the studies that included auditory working memory tests, few directly compared reading and listening working memory measures in relation to speech recognition in the same sample. It is important to compare the association between auditory and visual tests of verbal working memory and various measures of speech understanding in different listener groups before deciding on how specific test(s) could be used by rehabilitative audiologists. Of course, testing reading working memory rather than listening working memory to assess inter-individual differences in speech understanding would be a reasonable choice if reading and listening working memory tests yielded similar results, but assumptions about the modality-independence of working memory based on research in normal young listeners need to be confirmed in older adults and in listeners who have various degrees and types of hearing loss.

Mixed findings have been reported in a series of three recent Dutch studies examining the associations between measures of reading (Dutch version of the Daneman and Carpenter, 1980 test) and listening span (an auditory version of their reading span version presented in quiet) and a sentence-in-noise repetition task (Versfeld et al., 2000) in younger or middle-aged adults. In the first study (Koelewijn et al., 2012), middle-age listeners ($n = 32$; mean age = 51.3 years) with normal hearing were tested and a significant correlation between the reading and listening span measures (Pearson $r = 0.67$) was found; there also were significant correlations between reading span and sentence recognition thresholds in fluctuating and single-talker maskers (Pearson $r = -0.36$ to -0.50), but no significant correlations between listening span and the sentence-recognition thresholds. In another study using the same Dutch measures in younger adults ($n = 24$) with normal hearing (Zekveld et al., 2013), no significant correlations between the two span measures were found and neither span measure correlated significantly with the scores on the sentence-in-noise repetition task. However, in a third study (Besser et al., 2013) using the Dutch measures in younger listeners with normal hearing ($n = 42$) in two sessions (test–retest purposes), there was a significant correlation between the span measures administered in the two modalities (Pearson $r = 0.49$ in session 1 and $r = 0.60$ in session 2), but neither span measure correlated with speech-in-noise performance in either session. Taken together, these studies suggest that measures of reading and listening span in quiet are usually significantly correlated, but correlations between span measures and speech-in-noise thresholds for speech recognition are elusive for reading span and absent for listening span in quiet when younger or middle-aged adults with normal audiometric thresholds are tested. It is possible that little working memory resources are required by these listeners in these test conditions.

In contrast, studies comparing younger adults to older adults with normal or near-normal, hearing suggest that listening working memory span (LWMS) in quiet may be a

more informative measure than reading working memory span (RWMS). Pichora-Fuller et al. (1995) measured RWMS and LWMS in a group of younger listeners with normal hearing ($n = 16$) and in a group of older listeners with normal hearing through 3000 Hz ($n = 16$). The two tests followed the same protocol for determining working memory span. The reading measure used the same sentences as had been used in earlier studies by Daneman and Carpenter (1980, 1983). The listening measure used sentences from the Revised-Speech in Noise test (R-SPIN; Bilger, 1984) presented in quiet. Their results showed a significant correlation between the reading and LWMS measures for both the younger ($r = 0.56$) and older ($r = 0.71$) listener groups. Although an age-related difference in RWMS often is found (see Bopp and Verhaeghen, 2005 for a meta-analysis), in the study of Pichora-Fuller et al. (1995), both age groups had equivalent performance on the RWMS measure, perhaps because those in the older group were cognitively high-performing, well-educated, healthy older adults. Notably, despite the equivalent performance of the two age groups on the RWMS measure, younger adults had larger (better) LWMSs than the older adults and the older group performed worse on the LWMS test than on the RWMS test. The authors attributed this pattern of findings to age-related differences in supra-threshold auditory processing rather than to general modality-independent age-related differences in cognition, consistent with the domain-specific view of working memory. Similar results were found in a more recent study (Baldwin and Ash, 2011) in which RWMS and LWMS (in quiet) and speech-recognition threshold in quiet were tested in a group of younger ($n = 80$) and older ($n = 26$) adults with normal audiometric pure-tone thresholds through 8000 Hz. Specifically, the RWMS scores of the two age groups were similar, but the LWMS and speech recognition threshold results were significantly poorer for the older listeners compared to the younger listeners. A Pearson r correlation between RWMS and LWMS was not reported; however, a regression analysis showed that speech recognition thresholds, but not RWMS, predicted LWMS performance in older listeners, but not younger listeners. These studies comparing younger and older adults with normal or near-normal hearing (Pichora-Fuller et al., 1995; Baldwin and Ash, 2011) suggest that measuring LWMS may reveal age-related inter-individual differences relevant to listening performance on speech tests that are not revealed by measuring RWMS.

Another reason some studies may have been more successful than others in finding an association between measures of working memory and speech recognition is the selection of speech materials. Most studies have used various sentence-level materials in various listening conditions (e.g., quiet, noise, aided, etc.; see Akeroyd, 2008 and Besser et al., 2013 for reviews). Other studies have used phoneme or word-based materials (e.g., Akeroyd, 2008 for review; also see Humes and Floyd, 2005; Cervera et al., 2009; Baldwin and Ash, 2011; Smith et al., under review). A few studies have investigated associations between working memory and a range of speech materials in the same participants. For example, Humes and Floyd (2005) examined the associations among working memory (measured using a Simon-Says memory game (Pisoni and Cleary, 2004), presented in an auditory-only, visual-only, or auditory-visual condition) and two speech measures, a nonsense syllable test (City University of New York Nonsense Syllable Test [CUNY NST], Levitt and Resnick, 1978) and an open-set sentence recognition task (Connected Speech Test [CST], Cox et al., 1988), presented in unaided and aided conditions, in younger listeners with normal hearing ($n = 12$) and older listeners with hearing loss ($n = 24$; correlations were based on data for 22 of the 24 older listeners with hearing loss). Regardless of the modality, the Simon-Says task was not correlated with either speech measure in either condition in this study. In contrast, Cervera et al. (2009) did report significant correlations; specifically, they examined associations between two memory tests (serial recall and digit ordering) and two speech tests, a vowel-consonant-vowel (VCV) nonsense syllable repetition test (in quiet and in noise) and an open-set sentence recognition test (normal and fast speech rate) in 28 younger adult listeners with normal hearing and 27 older participants (mean age = 60 years) with mild, high-frequency hearing loss. The results showed that memory measures did not correlate significantly with the VCV materials, but a significant correlation emerged for both memory measures and fast-rate sentence recognition. These two studies illustrate the range of memory tests and speech materials used in listeners with and without hearing loss and across age groups (see also the reviews by Akeroyd, 2008; Besser et al., 2013).

In summary, a number of working memory and speech measures have been used to examine associations between working memory and speech understanding in adults with and without hearing loss. Discrepancies in findings may be attributable to the participants, the materials and the tasks used across the studies. We aimed to explore the associations between verbal working memory measures presented in the visual and auditory modalities and to determine if there would be modality-specific associations depending on the linguistic level of the materials (words, sentences, discourse), the nature of the task (simple repetition vs. comprehension) used to test speech understanding, and the age and hearing loss of the listener group. We hypothesized that there would be a significant correlation between LWMS and RWMS for all three groups, but that LWMS would be more strongly correlated than RWMS with speech measures, especially in the older listeners with hearing loss, when more linguistically complex materials were used and for the task involving comprehension rather than simple repetition of the speech materials.

Materials and Methods

Participants

Three listener groups participated ($n = 24$ per group)[1]. One group consisted of younger adults with normal hearing (YN; mean age = 23.5 years, $SD = 2.8$, range = 19–29; 7 male) who were recruited from the Johnson City, Tennessee community.

[1]These groups of participants readily were available and were chosen to enable comparison to the common participant groups in prior studies. In particular, the two older groups with hearing loss are important because they were drawn from the clinical population of interest and provide a contrast in terms of age while matching on audiometric thresholds. The younger group with normal hearing provide an

The other two groups were older adults with hearing loss and were Veterans recruited from the Mountain Home, Tennessee Veterans Affairs (VAs) Medical Center Audiology clinic. The 'young–old' group (YOHL) had a mean age of 66.3 years ($SD = 2.0$, range = 63–69; 24 male), and the 'older' group (OHL) had a mean age of 74.3 years ($SD = 3.2$, range = 70–80; 24 male). A one-way analysis of variance (ANOVA) confirmed a significant difference in age among the three listener groups, $F(2,71) = 2416.4$, $p < 0.001$. The average education level was 15.3 years ($SD = 2.1$, range = 12–20) for the YN listeners, 14.9 years ($SD = 2.7$, range = 12–20) for the YOHL listeners, and 13.9 years ($SD = 2.5$, range = 8–18) for the OHL listeners; a one-way ANOVA indicated no significant group difference in education level ($p > 0.05$). **Figure 1** illustrates the average audiogram of the test ear of the three listener groups (right ear of even-numbered participants and left ear of odd-numbered participants). A repeated-measures ANOVA for audiometric thresholds across frequency of the test ear (within-subjects factor) with hearing loss groups (YOHL and OHL) as between-subjects factors, revealed no significant main effects of group, nor was there a frequency by group interaction ($p > 0.05$), suggesting similar test-ear audiograms for the YOHL and OHL groups.

The inclusion criteria were as follows: ability to speak American English; adequate vision and ability for reading verified by reading aloud a few sentences from the informed consent document; ≥50% correct word recognition accuracy in quiet to avoid floor effects with the test materials; >21/30 on the Montreal Cognitive Assessment to rule out dementia (Nasreddine et al., 2005); and no comorbid health condition (e.g., conductive hearing loss, substance abuse, blindness, mental health disorder, etc.) that potentially would interfere with the study procedures as determined by an interview (younger adults) or medical records review (older adults). Although tinnitus is a potential comorbid condition that may interfere with working memory (e.g., Rossiter et al., 2006), a positive history of tinnitus was not used as an exclusionary criterion[2].

extreme contrast in terms of both age and hearing thresholds. We recognize that inclusion of a group of older listeners with normal hearing and a group of younger listeners with hearing loss would have offered a more ideal examination regarding the effects of age and hearing loss; however, older adults with normal hearing would not usually be seen in audiology clinics and the underlying mechanisms of hearing loss in younger adults are not the same as those of age-related hearing loss even though audiometric thresholds may similar.

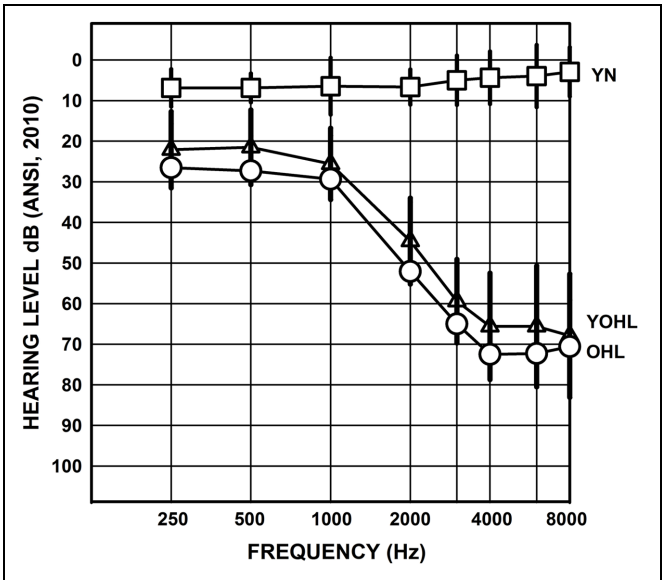


FIGURE 1 | Mean audiogram of the test ear for the younger listeners with normal hearing (YN; squares), young-old listeners with hearing loss (YOHL; triangles), and older listeners with hearing loss (OHL; circles). The error bars represent one standard deviation.

Materials

A battery of five memory measures (three auditory and two visual) and six auditory measures of speech understanding were administered to each participant. These measures were chosen because of their availability and prior use in research and clinic applications. The memory measures included free recall and working memory presented in both the auditory and visual domains. The tests of speech understanding used a continuum of materials that varied in linguistic complexity (word, sentence, or discourse level materials) and tasks (simple repetition or comprehension). All auditory test materials were pre-recorded and most were spoken by the same talker (VA female speaker #2) drawn from a corpus of materials recorded by Wilson et al. (2008).

Memory Tests

Reading Span (RS; Daneman and Carpenter, 1980)

The reading span test is a verbal working memory test administered in the visual domain using text. A total of 100 sentences are presented in five setsizes (2, 3, 4, 5, or 6 sentences per set) with five trials at each setsize. Thus, there are five 2-sentence trials (10 sentences); five 3-sentence trials (15 sentences); five 4-sentence trials (20 sentences), five 5-sentence trials (25 sentences), and five 6-sentence trials (30 sentences). The participant sees one sentence at a time (text via power point) and is asked to (1) read the sentence aloud, (2) make a judgment about whether or not each sentence makes sense (which serves to induce semantic processing of the entire sentence), and (3) at the end of a trial when prompted with a blank blue screen, the participant recalls the final word from each sentence in the trial in the order in which they were presented. The RS test was scored in terms of span size or the largest setsize for which the participant correctly recalls three out of five trials; however, partial credit is given for up to two out of three correctly recalled trials in the next highest set size.

[2]Tinnitus information was available only for the older listeners via a chart review. The majority of older listeners (40/48) reported experiencing tinnitus in some way as an adult (e.g., history of tinnitus would be positive even if they reported tinnitus occurring rarely, or only for a few minutes, etc.) and only 17.5% (7/40) reported that their tinnitus was bothersome in some way (i.e., can interfere with sleep at times, etc.). There were no differences on the two primary memory measures as a function of tinnitus being: (1) positive vs negative history, (2) constant vs intermittent, or (3) bothersome vs not bothersome. Future studies in this area should consider tinnitus for inclusion/exclusion purposes.

Word Auditory Recognition and Recall Measure (WARRM; Smith et al., under review)

The WARRM is an auditory working memory measure. The general procedures of the WARRM follow the RS test paradigm, but audio-recorded (VA female speaker #2) monosyllabic words following a standard carrier phrase "*You will cite . . .* " are used as the target items. As in the RS test paradigm, in the WARRM, there are 100 targets presented across five setsizes (2, 3, 4, 5, and 6 per set) with five trials being tested for each setsize. The participant is presented one item at a time and is asked to (1) repeat aloud the target word, (2) make a judgment about whether the first letter of the word is from the first half (A-M) or the second half (N-Z) of the alphabet (which serves to induce further processing of the word to be recalled), and (3) the participant recalls all of the target words in the trial in the order in which they were presented when prompted with 500-Hz, 500-ms tone at the end of a trial. The WARRM yields two scores, a word recognition accuracy score (percent correct), which served as one of the six speech measures in the current study (also described below), and a working memory span score.

Visual Free Recall (VFR; Adapted from Rabbitt, 1968, 1991)

This test uses a list of 15 words with each word presented individually on a plain white power-point slide with a 1-s inter-stimulus interval (ISI). After the series of words are presented, a yellow slide with the word 'RECALL' in black is used as the recall prompt. The participants are asked to write down as many words as they can recall from the list on a score sheet in 3 min. The test is scored by summing the number of correctly recalled words.

Auditory Free Recall (AFR; Adapted from Rabbitt, 1968, 1991; Park et al., 1996)

Analogous to the VFR test, a list has 15 audio-recorded (using VA female speaker #2) monosyllabic words presented individually with a 2-s ISI between words. Following the series of words, a 500-Hz, 500-ms prompting tone is presented to cue recall. There are no common words between the AFR and VFR measures.

Digit Span (DS)

A modified audio version of the Wechsler Adult Intelligence Scale (fourth edition, WAIS-IV, Wechsler, 2008) digit span (DS) subtest was used. Typically, the test is administered in a face-to-face interview format in which the examiner presents trials by live voice. A trial consists of a series of single digits spoken at a rate of one per second. The number of digits per trial increases during the test, with two trials for each span size, starting with a 2-DS size and terminating with a 9-DS size. Rather than the typical live voice test presentation method, to ensure a more standardized method of administration (e.g., consistent ISI, talker, and presentation level), the test was modified by using a series of monosyllabic digits (0 and 7 were replaced with monosyllabic digits) recorded by VA female talker #2, followed by a 500-Hz, 500-ms prompting tone. Otherwise, the general procedures of the DS test were maintained for the digit span forward (DSF), digit span backward (DSB), and digit span sequencing (DSS) subtests. For all subtests, the listener is presented with a series of digits, presented one at a time with a 1-s ISI, followed by the prompting tone. The response required from the listener varies with each subtest in that the listener is asked to recall the digits in the order in which they were presented (DSF), in the reverse order in which they were presented (DSB), or in the ascending numerical order in which they were presented (DSS). The subtests are scored by summing the number of correctly recalled trials.

Speech Understanding Tests

Word Recognition in Quiet (from the WARRM)

An overall percent correct word recognition score across the 100 WARRM test items was calculated. This score served to determine word-recognition abilities in quiet for the same items for which recall also was tested (see above).

Words-In-Noise Test with VA Female Speaker #2 (WIN#2)

The original Words-In-Noise test (Wilson, 2003; Wilson et al., 2003) has two, 35-word lists presented in a six-talker background. The words are from the Northwestern University Auditory Test No. 6 (NU-6, Tillman and Carhart, 1966). For each list (List 1 and List 2), five words are presented at seven SNRs from 24- to 0-dB in 4-dB decrements. The WIN#2 test was modified by replacing the original NU-6 words with the same words spoken by VA female speaker #2 with the carrier phrase "*You will cite*" instead of the original "*Say the word*" carrier phrase. In the current study, the WIN#2 is scored by calculating the 50% point threshold (dB S/N) using the Spearman-Kärber equation and averaged across both lists (Finney, 1952; Wilson et al., 1973).

Multi-Signal-to-Noise Ratio Revised Speech in Noise Test (Multi-SNR R-SPIN; Wilson et al., 2012)

A modified version of the Revised Speech in Noise Test (R-SPIN; Bilger, 1984) was used. In this version, two 50-sentence lists containing R-SPIN sentences (from Lists 3 and 4, original male talker) were distributed across 10 signal-to-noise ratios (SNR, S/N) from 23- to 4-dB in 3-dB decrements, with five sentences at each SNR. Across the two lists and at each SNR, five low-probability (LP) and the corresponding five high-probability (HP) sentences were used. The listener is asked to repeat aloud the final word in each sentence. The test is scored by calculating separate 50% points (dB S/N via the Spearman-Kärber equation) for the sentence-final target words in the LP and HP sentences across the list pair, and there also is a linguistic context score (difference in 50%-point between HP and LP scores).

Quick Speech-in-Noise Test (QuickSIN; Killion et al., 2004)

Lists 1 and 2, along with a practice List A, of the QuickSIN were used (Etymotic Research, 2001). Each QuickSIN list consists of six Institute of Electrical, and Electronics Engineers (1969) sentences that are presented in a multi-talker background. One sentence is presented at each of 6 SNRs that range from 25- to 0-dB in 5-dB decrements. Each sentence is scored based on correct recognition of five keywords (e.g., A white silk jacket goes with any shoes.). In the current study, this test was scored in terms of the 50%-point

(Spearman-Kärber) and an overall QuickSIN score was calculated by averaging the scores across Lists 1 and 2.

Veterans Affairs Sentence Test (VAST; Bell and Wilson, 2001)

The VAST sentences are constructed based on the Neighborhood Activation Model (Luce, 1986). Briefly, the monosyllabic words selected for the sentences are based on four lexical categories including: (1) sparse, or words that are unique or have few similar "neighbors," (2) dense, or words with many phonetic similarities with other words (i.e., many lexical neighbors), (3) low use, or words that are infrequently used in spoken language, and (4) high use, or words that are frequently used on spoken language. Using these categories, four combinations of sentences types based on word frequency (either low or high use) and neighborhood similarity (either sparse or dense) were used to construct the VAST sentence lists, which included (1) low use, sparse (LS), (2) low use, dense (LD), (3) high use, sparse (HS), and (4) high use, dense (HD). Each participant was administered one 20-item VAST list that consisted of items from each sentence type (LS, LD, HS, and HD). Each sentence contains three keywords, and accuracy is scored in percent correct for the keywords for each list (60 keywords per list).

Lectures, Interviews, and Spoken Narratives Test (LISN; Tye-Murray et al., 2008)

Three spoken narratives (about 3 min each) from this test were used; two test narratives (Narrative 6 about an individual's college experience, male talker; Narrative 7 about a store fire, male talker) along with a practice narrative (Narrative 10 about a grocery store robbery, female talker). The narratives were spoken by different talkers in a natural, conversational style. Participants listened to each narrative in its entirety and answered six multiple-choice comprehension questions, each with four response choice alternatives (pen/paper format). These questions asked about three different aspects of listening comprehension including: (1) information (i.e., recalling a specific detail in the narrative), (2) integration (i.e., the listener's ability to combine pieces of information), and (3) inferences (i.e., the listener's ability to infer implications from the narrative). There are two questions for each aspect of listening comprehension. An overall listening comprehension score along with a score for each question type was calculated for each list and averaged across lists as a percent correct score.

Procedures

The study was approved by the local research ethics committees (East Tennessee State University/VA Institutional Review Board and VA Research and Development Committee). All groups provided informed consent prior to testing. After consenting, a pure-tone audiogram was obtained for the test ear (odd-numbered participants received testing in the left ear and even-numbered participants in the right ear) for octave frequencies of 250–8000 Hz and the inter-octave frequencies of 3000 and 6000 Hz (American National Standards Institute [ANSI], 2010). The YOHL and OHL listeners were administered a 25-word NU-6 list to ensure they had adequate word recognition abilities ($>50\%$) in the test ear to complete the protocol. All groups received the MoCA to ensure that no participant had a positive screen for dementia).

All visually presented materials were administered in a quiet lab space while the participant was seated at a table. The RS and VFR tests were administered using a computer (Dell, Model Optiplex 780) and a 15-inch computer screen (Dell 1908FP). Participants wore their habitual corrective lenses during testing if needed for reading. The YOHL and OHL listeners either wore their hearing aids (if they owned them) or a pocket talker during MoCA administration (Dupuis et al., 2015) and when test instructions were given to ensure they could hear the instructions optimally.

All audio-record materials were presented from a compact disc (CD) that was calibrated and then played through a CD player (Sony, Model CDP-CE375) routed through an audiometer (Grason-Stadler, Model 61) to an insert earphone (Etymōtic, Model ER-3A) while the participant was seated in a double-walled sound-attenuating booth. The NU-6 words, WARRM, modified DS, AFR, VAST, and LISN were all presented in quiet at presentation levels of 62 dB HL for YN listeners, 72 dB HL for YOHL and OHL with pure-tone averages (PTA at 500, 1000, and 2000 Hz) $<$ 40 dB HL, and 82 dB HL for YOHL and OHL with PTAs 40–60 dB HL. The WIN#2 and multi-SNR R-SPIN were presented at 80 dB SPL (equivalent to 62 dB HL) for listeners with PTAs $<$ 40 dB HL, and 90 dB SPL (equivalent to 72 dB HL) for listeners with PTAs 40-60 dB HL, with the levels used for the WIN#2 and multi-SNR R-SPIN based on the level of the noise, which was held constant while the level of the speech was varied to yield the range of SNRs tested. The presentation level of the QuickSIN lists followed the administration manual and were presented at 70 dB HL for participants with PTAs $\leq$ 45 dB HL and at a dial level that was "loud, but OK" for participants with PTAs $\geq$ 50 dB HL.

All listener groups completed the testing in two sessions. The tests for the experimental protocol were sequenced so that the tests were balanced across sessions to avoid fatigue and order effects. Session One lasted ∼80–90 min for each listener group. After consenting and testing for inclusion/exclusion criteria in Session One, all groups then were administered the RS and WARRM tests; the order of the tests was counterbalanced. The RS and the WARRM were grouped together because of similarities in their testing procedures. A 10-min break was required between these two working memory tests for the older groups, whose testing for Session One ended after the RS and WARRM testing was completed. For the younger listeners, there was a 10-min break required after the RS and WARRM testing, followed by the WIN#2 and the QuickSIN tests, with these tests counterbalanced across participants. Session Two lasted ∼60 min for the YN listeners and 90 min for older listeners. For Session Two, the session was divided into two halves, with one half of the session focusing on speech understanding testing and the other half of the session focusing on memory testing. The session halves were counterbalanced across participants and a 10-min break was required between the halves. For all groups, the memory testing half of Session Two included the DS, AFR and VFR measures. The DS and the VFR tests were administered in a

counterbalanced order, either first or last, with the AFR always being administered between them. The AFR and VFR tests were administered consecutively because of the similarities in the test procedure. The VFR test was either administered first or last in the session to minimize changes in test locations (either the sound booth or computer location) within the session half.

For the younger listeners, in the speech understanding testing half of Session Two, the LISN and VAST tests were counterbalanced, with the multi-SNR R-SPIN test always administered in between them because it was considered to be less demanding than the LISN and VAST tests. For the older listeners, in the speech understanding testing half of the Session Two, participants were administered the LISN, VAST, QuickSIN, WIN#2, and multi-SNR R-SPIN tests; the LISN or VAST were administered first or third (counterbalanced across participants) and the QuickSIN, WIN#2 or multi-SNR R-SPIN test were randomly assigned as the second, fourth, or fifth tests. The rationale for this ordering of tests was to administer a more demanding test followed by one that was less demanding to avoid fatigue for the older listeners. Because multiple lists were administered for a given speech understanding test, the list order of the speech tests also was counterbalanced to avoid order/list effects. For the QuickSIN and LISN tests only, a practice list was administered prior to the experimental lists. The four VAST lists were assigned randomly to each participant. The participants were encouraged to take additional breaks during testing as needed and were remunerated $20 per hour.

Results

Several measures were administered to three groups of participants (YN, YOHL, OHL) to assess their cognitive and speech understanding abilities. Descriptive results and group differences on each measure were calculated. Correlational analyses were performed to examine the associations between reading and LWMS. An ANOVA was conducted to examine the effect of test modality on working memory span. Finally, the contributions of memory to performance on various speech understanding measures were evaluated using correlational analyses. All data were analyzed with statistical software (International Business Machines Statistical Package for the Social Sciences, Version 22.0) and all analyses (ANOVAs and *post hoc* analyses) were adjusted (Bonferroni) to account for multiple comparisons.

In **Table 1**, the mean results for seven memory measures are listed for each group. For each variable, a separate one-way ANOVA was conducted to evaluate group differences and those results also are presented in the table. The ANOVAs revealed significant differences among the results for the groups on all memory measures. In all cases where there was a significant group difference, *post hoc* analyses showed that the younger listeners performed best, and the two groups of older listeners had similar performance that was significantly poorer than that of the younger listeners.

For each listener group, correlations were computed to explore the associations among the memory measures (only $ps < 0.007$ were considered to be significant). For the listeners in the YN group, Pearson r correlations were significant between AFR and DSB ($r = 0.61$, $p = 0.002$) and between AFR and DSS ($r = 0.59$, $p = 0.002$). No significant correlations were found among the memory measures for the YOHL listeners. For the OHL listeners, Pearson r correlations were significant between WARRM span and DSS ($r = 0.55$, $p = 0.006$); WARRM span and VFR ($r = 0.55$, $p = 0.005$); and DSB and DSF ($r = 0.63$, $p = 0.001$). Correlations between the RS and WARRM span will be presented later as they address a distinct aim of the study.

TABLE 1 | The mean performance (and one standard deviation) on the seven memory measures by the three listener groups.

	YN		YOHL		OHL				
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*df*	*p*
Reading span	2.5	0.9	1.9	0.3	1.9	0.4	8.4	2, 69	**0.001**
WARRM span	4.4	1.1	3.0	0.6	2.9	0.8	23.7	2, 69	**0.000**
Visual free recall	8.0	1.7	4.8	1.5	4.2	1.7	37.1	2, 69	**0.000**
Auditory free recall	8.3	2.2	4.8	1.8	3.5	1.4	45.0	2, 69	**0.000**
Digit span									
Forward	10.3	2.2	9.1	1.8	9.1	2.4	2.4	2, 69	0.102
Backward	8.8	1.9	7.7	1.6	7.3	2.5	3.2	2, 69	0.047
Sequencing	10.1	1.8	9.4	1.5	8.0	1.9	3.2	2, 69	**0.000**

The results from separate one-way analyses of variances also are listed. YN, younger listeners with normal hearing; YOHL, young–old listeners with hearing loss; OHL, older listeners with hearing loss; WARRM, Word Auditory Recognition and Recall Measure. Shown in bold are p values < 0.007 which were considered to be significant after Bonferroni corrections were applied.

Table 2 lists the mean performance for each listener group from the six speech tests and subtests if applicable. The results of the one-way ANOVAs to evaluate group differences on the speech measures also are presented in **Table 2** (p values < 0.003 were considered to be significant). The ANOVAs revealed a significant difference among groups for each speech understanding measure except for the measures from the LISN test and the Use of Context measure from the multi-SNR R-SPIN test. For WARRM recognition, the Low Probability measure from the multi-SNR R-SPIN test, and the HS, HD and LS measures from the VAST test, the younger group performed the best, followed by the two older groups who performed similarly. A different pattern emerged for the WIN#2, the High Probability measure from the multi-SNR R-SPIN test, the QuickSIN, and the LD measure from the VAST test, with all three groups performing significantly differently from each other; the YN group performed best, followed by the YOHL group, with the OHL group performing worst.

For each listener group, correlations were computed to explore the associations among the speech understanding measures (only $ps < 0.003$ were considered to be significant). The significant correlations for the YOHL (below the diagonal) and the OHL (above the diagonal) listeners are listed in **Table 3**. Note the WARRM in **Table 3** refers to the word recognition score. The correlations for both hearing loss listener groups were mostly non-significant, with moderate to strong correlations for those correlations that were significant.

TABLE 2 | The mean performance (and one standard deviation) on the speech understanding measures by the three listener groups.

	YN		YOHL		OHL				
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*df*	*p*
WARRM recognition (%)	*99.0*	*1.1*	79.7	9.2	76.2	13.5	40.5	2, 69	**0.000**
WIN#2 (dB S/N)	*3.0*	*1.0*	*14.0*	*3.3*	*16.0*	*3.3*	155.4	2, 69	**0.000**
multi-SNR R-SPIN									
Low-Probability (dB S/N)	*5.4*	*1.4*	9.7	2.8	10.9	3.1	30.5	2, 69	**0.000**
High-Probability (dB S/N)	*1.5*	*1.2*	*4.2*	*2.0*	*5.6*	*2.3*	29.6	2, 69	**0.000**
Use of Context (dB)	3.9	1.4	5.5	1.8	5.2	1.9	6.3	2, 69	0.003
QuickSIN (dB S/N)	*3.2*	*1.3*	*8.9*	*3.5*	*11.2*	*3.8*	43.8	2, 69	**0.000**
VAST (%)									
Low Use, Sparse	*98.5*	*2.2*	90.9	5.4	87.4	8.2	22.5	2, 69	**0.000**
Low Use, Dense	*98.4*	*2.1*	*92.4*	*4.9*	*88.0*	*7.9*	*21.9*	2, 69	**0.000**
High Use, Sparse	*99.4*	*1.2*	96.0	3.8	94.1	4.5	14.8	2, 69	**0.000**
High Use, Dense	*99.2*	*1.0*	93.5	3.9	92.8	5.3	20.2	2, 69	**0.000**
LISN (%)									
Overall	76.7	14.5	72.9	15.4	62.8	19.0	4.6	2, 69	0.014
Information	80.2	18.0	74.0	21.5	61.5	23.3	4.9	2, 69	0.010
Integration	71.9	17.0	78.1	17.0	75.0	23.3	0.6	2, 69	0.537
Inferences	78.1	25.9	66.7	21.7	52.1	27.5	6.5	2, 69	0.003

The results from separate one-way analyses of variances also are listed. YN, younger listeners with normal hearing; YOHL, young–old listeners with hearing loss; OHL, older listeners with hearing loss. WARRM, Word Auditory Recognition and Recall Measure (recognition score); WIN#2, Words-In-Noise Test Number 2; QuickSIN, Quick Speech in Noise test; multi-SNR R-SPIN, multi signal-to-noise ratio Revised Speech in Noise test; VAST, Veterans Affairs Sentence Test; LISN, Lectures, Interviews and Spoken Narratives test. Only p values < 0.003 were considered significant and are bolded. Italics are used to indicate the two patterns of results, either that the results of the younger group differed from those of the two older groups which did not differ from each other (only results of the younger group are italicized) or that all three groups differed significantly from each other (results for all three groups are italicized).

TABLE 3 | The Pearson *r* correlations among the speech measures for the YOHL group (below the diagonal) and for the OHL group (above the diagonal).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
(1) WARRM		−0.67	.	.	.	−0.76	0.83	0.66	.	.	.	.	.	.
(2) WIN#2	−0.77		0.60	0.69	.	0.68	−0.67	.	.	.	.		.	
(3) LP	.	0.73		0.79	0.69	.	.	.	.	.	.	.	.	.
(4) HP	.	0.64	0.76		.	.	.	.	.	.	.	.	.	.
(5) Context	.	.	0.71	.		.	.	.	.	.	.		.	
(6) QuickSIN	−0.71	0.74	.		.		−0.63	.	.	−0.59	.		.	
(7) LS	0.67	.	.	.	.	−0.66		0.77	.	.	.	.	.	.
(8) LD	.	.	.		.	−0.67	.		0.59	0.74	.		.	
(9) HS	.	−0.61	.		.	.	0.67	.		0.59	.		.	
(10) HD	0.68	.	.	.	.	.	0.65	.	.		.	.	.	.
(11) LISN	.	.	.	.	.	.	.	.	.	.		0.72	0.80	0.80
(12) Info.	.	.	.	.	.	.	.	.	.		0.87		.	.
(13) Integ.	.	.	.	.	.	.	.	.	.		0.65	.		.
(14) Infer.	.	.	.	.	.	.	.	.	.		0.76	.	.	

YOHL, young–old listeners with hearing loss; OHL, older listeners with hearing loss. WARRM, Word Auditory Recognition and Recall Measure (recognition score); WIN#2, Words-In-Noise Test Number 2; LP, low probability multi signal-to-noise ratio Revised Speech in Noise test (multi-SNR R-SPIN); HP, high probability multi-SNR R-SPIN; Context, multi-SNR R-SPIN Use of Context; QuickSIN, Quick Speech in Noise test; LS, low usage, spare Veterans Affairs Sentence Test (VAST); LD, low usage, dense VAST; HS, high usage, sparse VAST; HD, high usage, dense VAST; LISN, Lectures, Interviews and Spoken Narratives Test overall score; Info., information score on LISN; Integ., integration score on LISN; and Infer., Inference score on LISN.

For the YN listeners whose results are not listed in **Table 3**, significant Pearson r correlations were observed between the QuickSIN and the VAST LS ($r = -0.66$, $p < 0.001$). For the LISN test, the overall score was significantly correlated with the LISN information score ($r = 0.86$, $p < 0.001$) and LISN inference score ($r = 0.78$, $p < 0.001$). For the multi-SNR R-SPIN test, the Low Probability measure was significantly correlated with Use of Context measure ($r = 0.64$, $p < 0.001$). No other significant correlations among the speech measures for YN listeners were found.

The results obtained for the RS (visual) and WARRM (auditory) working memory tests were compared to evaluate differences due to test modality. **Figure 2** illustrates the mean performance on the RS and WARRM tests for each listener

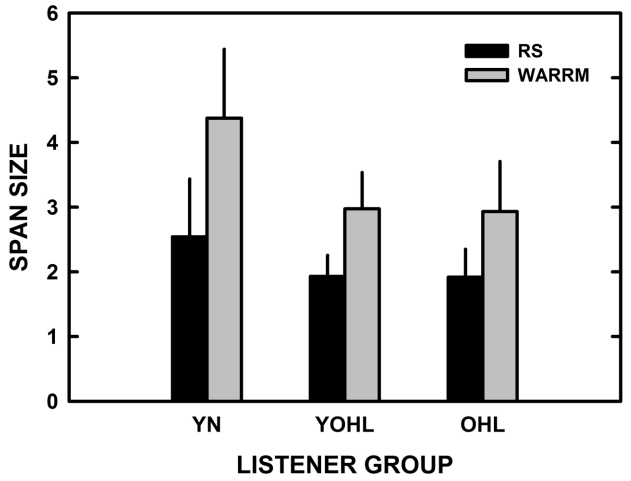


FIGURE 2 | The mean reading span (RS; black) and Word Auditory Recognition and Recall Measure (WARRM) span size scores (gray) are plotted as a function of listener group. The error bars represent one standard deviation. YN, young–old listeners with hearing loss; YOHL, young–old listeners with hearing loss; and OHL, older listeners with hearing loss.

group. A repeated measures ANOVA with group as the between-subjects variable (YN, YOHL, and OHL) was performed using span scores to compare test modalities (visual with the RS and auditory with the WARRM) as the within-subjects variable. The results showed a main effect of modality, $F(1,69) = 172.5$, $p < 0.001$, $\eta_p^2 = 0.71$, a main effect of group, $F(2,69) = 23.9$, $p < 0.001$, $\eta_p^2 = 0.41$, and a group by modality interaction, $F(2,69) = 7.4$, $p < 0.001$, $\eta_p^2 = 0.18$. *Post hoc* analyses showed that for the main effect of group (collapsed across RS and WARRM), the younger group performed best, followed by the two older groups, who had similar performance. For the main effect of modality (collapsed across group), performance was better on the WARRM span auditory test compared to the visual RS test. For the group by modality interaction, all groups performed better on the WARRM span (auditory) test relative to the RS (visual) test, but the difference between performances on these measures was larger for the younger listeners with normal hearing compared to the older listener groups who had similar differences in performance between the span measures.

For each group separately and for all participants combined, Pearson r correlations were conducted to examine the associations between RS and WARRM span scores (see **Figure 3**). For all groups, the correlation was $r = 0.52$, $p < 0.001$ (significant at the 0.01 level, two-tailed). When correlations were computed for each group, the only significant correlation was for the OHL group ($r = 0.55$, $p = 0.006$).

For each group, separate correlation analyses (controlling for high-frequency pure-tone average of 1000, 2000, and 4000 Hz) were conducted to examine the associations between the RS and WARRM span measures and each speech understanding measure. The only significant correlation found for the YOHL group was between the RS and WIN#2 scores ($r = 0.49$, $p = 0.02$; see **Figure 4**). For the YN listeners, WARRM span was significantly correlated with the QuickSIN ($r = -0.48$, $p = 0.02$) and RS was correlated with LISN information ($r = 0.47$, $p = 0.02$; see **Figure 5**). Aside from the few significant correlations, the general lack of significant correlations did not support our hypotheses that working memory would be correlated with results on tests of speech understanding and that the correlations would strengthen as the linguistic complexity of speech materials increased, particularly for OHL listeners. In fact, there were no significant correlations between working memory and speech understanding measures for the OHL listeners.

For the current study, we selected speech measures with the presumption that there would be increasing demands on working memory as linguistic complexity increased from words, to sentences, and then discourse. We expected that RS and WARRM would be significantly correlated with performance on tests of speech understanding, but that the strengths of those correlations would depend on the linguistic properties of the speech materials. In addition, we expected the strength of the correlations to be stronger for WARRM than RS depending on the auditory abilities of the participants. Because our hypotheses were not supported by the correlational analyses, we conducted a factor analysis to examine further the relations among the measures of memory and speech understanding. To this end, a principal components factor analysis using varimax rotation was conducted. Data from all participants ($n = 72$) were included. All speech understanding measures and memory measures, along with age and degree of hearing loss (determined by the pure-tone average [PTA] of 500, 1000, and 2000 Hz), were inputted into the analysis. The results revealed a five-factor solution that explained 76.9% of the variance (**Table 4** shows factor loading values > 0.60 for all five factors). The scree plot, however, suggested that the first three factors may be the most appropriate components to include in the solution. In general, as can be seen in the table, the majority of the speech understanding measures, along with age and PTA (which typically are correlated with speech measures), loaded on Factor 1. The majority of the memory measures loaded on Factor 2. The LISN (sub)tests loaded on Factor 3 and the Use of Context score from the multi-SNR R-SPIN loaded on Factor 4. The DSF loaded on Factor 5. These results suggest that there is a similarity amongst age, PTA and the speech understanding measures when the speech understanding task is simply to repeat words or sentences, whereas the speech understanding measures involving the comprehension of discourse or the use of semantic context are separate factors. Importantly, the majority of the memory measures were distinct from both kinds of speech understanding measures, and also the more basic and less cognitively demanding DSF memory measure.

Discussion

The main study aim was to examine the effect of presentation modality (auditory or visual) on verbal working memory measures in different listener groups. As expected, there was a significant effect of group, with the YN listeners outperforming the YOHL and OHL listeners on verbal working memory

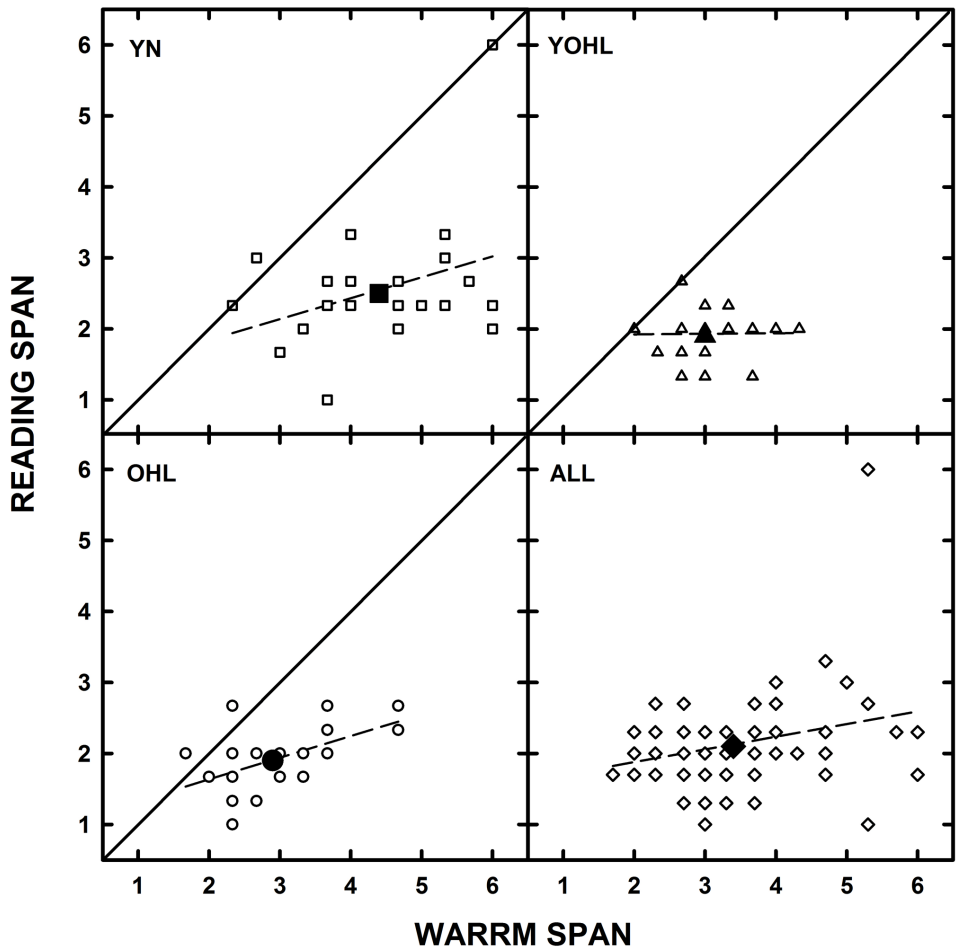


FIGURE 3 | The individual datum points (open symbols) for reading span are plotted as a function of Word Auditory Recognition and Recall Measure (WARRM) span scores for the younger listeners with normal hearing (YN; squares), young-old listeners with hearing loss (YOHL; triangles), older listeners with hearing loss (OHL; circles), and all participants in each panel respectively. The large-filled symbols represent the group mean data. The solid line represents equal performance and the dashed line represents the linear regression through the datum points.

measures tested in both modalities. Previous studies have demonstrated such age effects on working memory measures, in particular, in studies using the reading span measure (e.g., see Bopp and Verhaeghen, 2005 for a meta-analysis). Little data exists for the newly developed WARRM measure; however, previous data comparing 48 YN listeners with normal hearing to 48 older listeners with normal to near-normal hearing (ONH) and 48 older listeners with hearing loss revealed significant differences in mean WARRM spans suggesting that age affects performance on this measure (4.7 vs. 3.9 for the YN and ONH groups, respectively) and that hearing loss also affects performance (3.9 and 3.6 for the ONH and the OHL groups, respectively (Smith et al., under review).

The current results indicate that, for all listener groups, WARRM span was significantly higher and more variable than RS. There are a number of possible explanations for the difference in span size between the WARRM and RS tests. First, working memory span measures have been shown to be sensitive to the complexity of linguistic processing required for comprehending sentences (Waters and Caplan, 1996). Both the RS and WARRM measures use sentence-length stimuli, but the RS stimuli are a set of unique sentences, whereas the WARRM stimuli are monosyllabic words following a standard carrier phrase. Thus, because the WARRM stimuli are simpler and require less linguistic processing compared to the RS sentences, it would be expected that participants should be able to store more WARRM target words than RS target words. Second, for the RS measure, participants were asked to read aloud each sentence as they progressed through the recall set, thereby reducing the opportunity to rehearse the previous final words in the trial. In contrast, for the WARRM measure, each target word was presented following the same carrier phrase ('*You will cite*') and only the target word was repeated. Thus, even though the ISI between individual words on the WARRM was short (3 s) and was intended to leave time only for repetition of the target word and the linguistic judgment task, participants may have had more opportunity to rehearse the target words in the ISI or during the carrier phrase. Third, serial recall can be affected

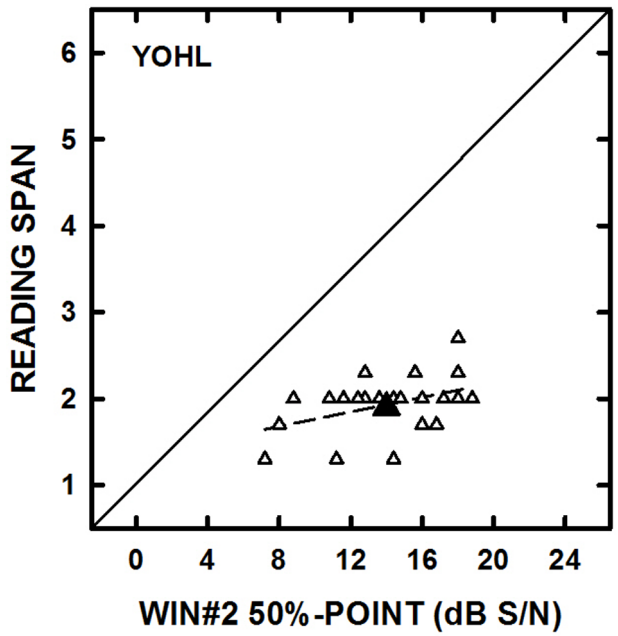


FIGURE 4 | The individual datum points (open symbols) for reading span are plotted as a function of performance on the Words-In-Noise #2 Test (WIN#2) for the young–old listeners with hearing loss (YOHL). The large-filled symbols represent the group mean data. The solid line represents equal performance and the dashed line represents the linear regression through the datum points.

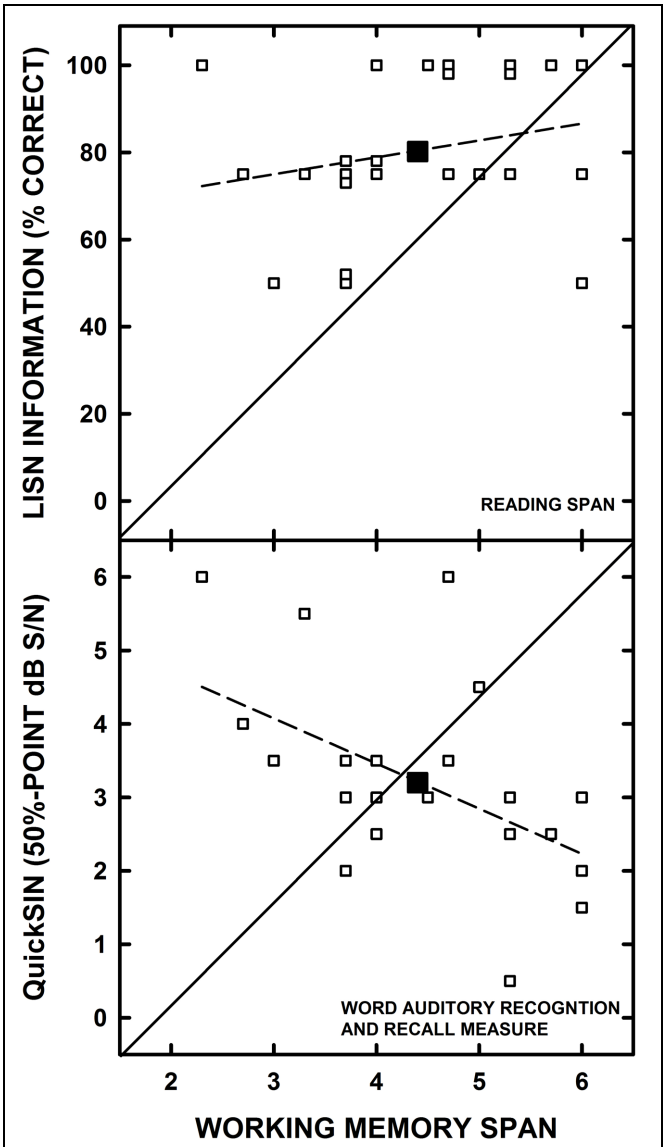


FIGURE 5 | The individual datum points (open symbols) for Lectures, Interviews and Spoken Narratives (LISN) Information scores are plotted as a function of Reading Span (top) and Quick Speech-in-Noise test scores are plotted as a function of Word Auditory Recognition and Recall Measure (WARRM) span scores are (bottom) for the younger listeners with normal hearing. For graphical clarity, the data were jittered slightly to offset overlapping datum points. The large-filled symbols represent the group mean data. The solid line represents equal performance and the dashed line represents the linear regression through the datum points.

by word length such that monosyllabic word sequences are recalled more accurately than are multi-syllable word sequences (e.g., Baddeley et al., 1975), possibly because of differences due to word length in rehearsal opportunity or forgetting during the recall response period (Baddeley, 2003). The final words to be recalled in the RS test included both monosyllabic and multi-syllabic words and the inclusion of multi-syllabic words may have resulted in more forgetting on the RS than in the WARRM test. In short, linguistic differences between the RS and WARRM stimuli may have differentially affected processing requirements, opportunities for rehearsal and propensity for forgetting, resulting in better performance in the WARRM span relative to the RS for all listener groups. It seems unlikely, however, that individual differences in linguistic abilities would have resulted in greater variability on the linguistically easier WARRM test compared to the more linguistically difficult RS test. Rather, less variability should have been observed on the easier WARRM test than on the harder RS test if linguistic processing were the explanation for inter-test differences.

A significant interaction between verbal working memory test modality and group was found. The interaction emerged because the difference between the two working memory measures was larger, almost twice as large for the YN listeners (1.9) relative to the two older listener groups (1.0 and 1.1, respectively; see **Table 1** and **Figure 2**). For the RS test, the differences in spans between the groups were small (by ∼0.5 span size), but the pattern of differences between groups did demonstrate the typical age effects. For the WARRM, the effect of age also was observed; however, there were larger group differences on the WARRM test (by about 1.5 span units of difference between YN and YOHL/OHL groups) compared to the RS test. The YN listeners have normal pure-tone thresholds and presumably better auditory processing relative to the two older listener groups who have hearing loss. It is likely that the higher WARRM spans for the YN listeners could be attributed to their relative ease

TABLE 4 | The factor loading values (sorted by strength) and the percent variance explained for each factor resulting from the factor analysis results are listed.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
WARRM Recognition	0.91				
VAST LS	0.86				
QuickSIN	−0.85				
WIN#2	−0.84				
VAST HS	0.83				
VAST LD	0.83				
VAST HD	0.81				
multi-SNR R-SPIN HP	−0.77				
Pure-tone average	−0.73				
Age	−0.70				
multi-SNR R-SPIN LP	−0.69				
AFR	0.64	0.60			
RS		0.76			
WARRM span		0.66			
VFR		0.64			
LISN overall			0.96		
LISN information			0.76		
LISN integration			0.73		
LISN inference			0.69		
multi-SNR R-SPIN context			0.89		
DSF					0.90
Percent variance	35.4	15.2	13.0	6.6	6.6

in hearing the WARRM stimuli compared to the older listeners with hearing loss. Accordingly, the difference between the two working memory measures within groups was largest for the YN compared to the other two groups, possibly reflecting differences in age and auditory processing abilities among the groups. It also seems reasonable that individual differences in auditory processing abilities might explain the greater variability observed in the results on the WARRM test than in the results on the RS test.

There was a significant moderate correlation ($r = 0.55$) between the RS and WARRM span measures for the OHL group only. Previous studies have found moderate correlations between LWMS and RWMS measures for younger (Pichora-Fuller et al., 1995; Besser et al., 2013), middle-age (Koelewijn et al., 2012), and older listeners with normal hearing (Pichora-Fuller et al., 1995). Although the current study did not demonstrate such correlations for YN and YOHL listeners, the results provide evidence that listening and reading span measures are moderately associated in older listeners with hearing loss. Furthermore, as can be seen in **Figure 3**, there is more variability in the individual datum points for WARRM span (abscissa) relative to RS (ordinate). Thus, the small range in performance on the RS likely contributed to a lack of a significant correlation between the measures for the YN and YOHL listeners. For researchers or clinicians interested in examining inter-individual differences in verbal working memory and how those differences relate to individual differences in speech understanding, given the greater range in performance on the WARRM test relative to the RS test, the WARRM may be a better metric to capture individual differences in verbal working memory across a range of listener groups.

The second aim of the present study was to examine the extent to which verbal working memory (RS or WARRM span) is associated with various measures of speech understanding for the different listener groups. Our hypothesis was that working memory would become more strongly correlated as the level of linguistic complexity of the materials increased (from word to sentence to discourse) and as the task shifted from simple repetition to comprehension. We also expected that WARRM span would be more strongly correlated than RS with measures of speech understanding, especially as linguistic complexity increased and especially for older adults with hearing loss. Contrary to our prediction, more significant correlations were found for YN listeners than the other groups, but the strength of the correlations did not change as a function of linguistic complexity or modality of the working memory measure. The observation of more significant correlations for the YN group may have arisen because their performance was not affected by hearing loss. Previous research has suggested that working memory emerges as a small, but significant factor explaining speech understanding, particularly speech-in-noise performance, only after audibility is accounted for, either by manipulation of the presentation level or through amplification, but that without correction for hearing loss the variance due to working memory is dominated by measures of hearing loss (Akeroyd, 2008; Houtgast and Festen, 2008; Humes et al., 2013). In the current study, the level of presentation of the speech stimuli was selected based on the hearing level of the participant; however, even with this correction for audibility, some high-frequency speech components may not have been fully audible for the YOHL and OHL listeners (see Humes, 2007 and Smith et al., 2012), whereas the YN group did not require any correction because they had normal hearing. Thus, we conclude that these results overall did not provide compelling evidence to support our hypotheses that there would be significant associations between measures of working memory and speech understanding. One reason for the lack of correlations may be that the current study was underpowered with 24 participants per group. Future work should test a larger sample size. Another reason may be because the entire speech signal was not fully audible in the older groups, thereby preventing the contribution of working memory to speech understanding from being fully realized in those listeners.

In light of the absence of significant correlations between measures of working memory and speech understanding, the factor analysis was performed to determine if indeed the speech measures were distinct and if the memory measures overlapped with the measures of speech understanding. The factor analysis indicated that the LISN test of discourse comprehension was unique relative to the other measures of speech understanding, but that the remaining measures of speech understanding based on a simple repetition task were not distinguishable enough to load on separate factors. In essence, whether word-level or sentence-level materials were used, the measures that loaded on Factor 1 employed a simple immediate word repetition task. For

example, for the QuickSIN and VAST tests, the task of the listener is to repeat the entire sentence, with the sentence being scored in terms of the number of keywords that are correctly repeated, whether or not the sentence that is repeated makes sense. For the multi-SNR R-SPIN, the whole sentence is presented, but the task of the listener is to repeat only the sentence-final word. Taken together, the findings that memory, repetition and comprehension measures were not correlated with each other and that they were not overlapping factors in the factor analysis, suggests that these factors are distinct and may depend as much if not more on task than on the linguistic nature of the test materials.

Another issue to consider is the ecological validity of using word recognition and comprehension measures as surrogates for everyday conversations. It could be that associations between memory and speech understanding measures would be significant if a more ecologically relevant measure of speech understanding, such as conversational fluency, were used rather than the relatively artificial and passive listening measures used in the current study. Additionally, in the present study, the measures used a mixture of materials spoken by different talkers and presented in quiet or in different types of babble. Future research examining the effects of linguistic complexity and task demands on the association between working memory and speech understanding should consider using a range of speech materials with the same talker in quiet and with consistent competing noise(s) to ensure that participants receive all levels of materials in all conditions with better control over the acoustic properties of the test materials. In addition, the effects of age and hearing loss may be better elucidated if groups of both younger and older adults with matched degrees of hearing thresholds (normal and with hearing loss) were used or if auditory performance was matched on the basis of other non-speech auditory measures of supra-threshold processing.

Conclusion

In summary, the data showed that all participants had better performance with the auditory WARRM test than with the visual RS test, most likely because the WARRM sentences were linguistically simpler and demanded less processing compared to the sentences used in the RS test. In addition, variability in verbal working memory was observed when participants were tested with the auditory WARRM test than with the visual RS test, most likely because the WARRM test was more sensitivity to individual differences in auditory processing. Furthermore, the findings did not provide overwhelming evidence that working memory is associated with various measures of speech understanding in any of these listener groups, regardless of age or hearing status. Instead, the findings suggest that measures of memory, word recognition and discourse comprehension tap distinct abilities that may be related to everyday listening and that these abilities should be measured separately. Future studies should use more consistent materials and methodological approaches to elucidate a better understanding regarding the possible associations between inter-individual differences in working memory and speech understanding in more ecologically relevant conditions.

Author Contributions

SS and MP-F both made contributions to the design of the work. SS oversaw data collection and analyzed the data. SS and MP-F both made contributions to interpretation of the work, drafting and revising the manuscript for important intellectual content, approving the final version to be published, and are accountable for all aspects of the work.

Acknowledgments

This work was supported by the Department of Veterans Affairs (VA), Veterans Health Administration, Office of Research and Development, Rehabilitation Research and Development (RR&D) Service, Washington D.C. by the RR&D Auditory Vestibular Research Enhancement Award Program (REAP; #C4339F) to the first author. The authors acknowledge Sam Hester, Kelsey King, Emerald Lauzon, and Devon Shock with their efforts with data collection.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

American National Standards Institute [ANSI] (2010). *American National Standard Specification for Audiometers* (*ANSI S3.6-2010*). New York, NY: Author.

Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x

Baddeley, A. (2003). Working memory and language: an overview. *J. Commun. Disord.* 36, 189–208. doi: 10.1016/S0021-9924(03)0 0019-4

Baddeley, A. D., Thomson, N., and Buchanan, M. (1975). Word length and the structure of short-term memory. *J. Verbal Learn. Verbal Behav.* 14, 575–589. doi: 10.1016/S0022-5371(75)80045-4

Baldwin, C. L., and Ash, I. K. (2011). Impact of sensory acuity on auditory working memory span in young and older adults. *Psychol. Aging* 26, 85–91. doi: 10.1037/a0020360

Bell, T. S., and Wilson, R. H. (2001). Sentence recognition materials based on frequency of word use and lexical confusability. *J. Am. Acad. Audiol.* 12, 514–522.

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459

Bilger, R. C. (1984). "Speech recognition test development," in *Speech Recognition by the Hearing Impaired. ASHA Reports 14*, ed. E. Elkins (Rockville, MD: American Speech-Language-Hearing Association).

Bocca, E., and Calearo, C. (1963). "Central hearing processes," in *Modern Hearing Developments in Audiology*, ed. J. Jerger (New York, NY: Academic Press), 337–370.

Bopp, K. L., and Verhaeghen, P. (2005). Aging and verbal memory span: a meta-analysis. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 60, P223–P233. doi: 10.1093/geronb/60.5.p223

Cacace, A. T., and McFarland, T. J. (2013). Factors influencing tests of auditory processing: a perspective on current issues, and relevant concerns. *J. Am. Acad. Audiol.* 24, 572–589. doi: 10.3766/jaaa.24.7.6

Cervera, T. C., Soler, M. J., Dasi, C., and Ruiz, J. C. (2009). Speech recognition and working memory capacity in young-elderly listeners: effects of hearing sensitivity. *Can. J. Exp. Psychol.* 63, 216–226. doi: 10.1037/a0014321

Committee on Hearing, Bioacoustics, and Biomechanics [CHABA] (1988). Speech understanding and aging. working group on speech understanding and aging. Committee on hearing, bioacoustics, and biomechanics, commission on behavioral and social sciences and education, national research council. *J. Acoust. Soc. Am.* 83, 859–895.

Cox, R. M., Alexander, G. C., Gilmore, C., and Pusakulich, K. M. (1988). Use of the Connected Speech Test (CST) with hearing-impaired listeners. *Ear Hear.* 9, 198–207. doi: 10.1097/00003446-198808000-00005

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Daneman, M., and Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 561–584. doi: 10.1037/0278-7393.9.4.561

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546

Dupuis, K., Pichora-Fuller, M. K., Chasteen, A. L., Marchuk, V., Singh, G., and Smith, S. L. (2015). Effects of hearing and vision impairments on the Montreal Cognitive Assessment. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 22, 413–437. doi: 10.1080/13825585.2014.968084

Etymotic Research (2001). *Quick Speech-in-Noise Test* [*Audio CD*]. Elk Grove Village, IL: Author.

Finney, D. J. (1952). *Statistical Method in Biological Assay*. London: Griffin.

Foo, C., Rudner, M., Rönnberg, J., and Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *J. Am. Acad. Audiol.* 18, 618–631. doi: 10.3766/jaaa.18.7.8

Frisina, D. R., and Frisina, R. D. (1997). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear. Res.* 106, 95–104. doi: 10.1016/S0378-5955(97)00006-3

Gates, G. A., and Mills, J. H. (2005). Presbycusis. *Lancet* 366, 1111–1120. doi: 10.1016/S0140-6736(05)67423-5

Houtgast, T., and Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *Int. J. Audiol.* 47, 287–295. doi: 10.1080/14992020802127109

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *J. Am. Acad. Audiol.* 18, 590–603. doi: 10.3766/jaaa.18.7.6

Humes, L., and Dubno, J. (2010). "Factors affecting speech understanding in older adults," in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (New York, NY: Springer), 211–257.

Humes, L. E., and Floyd, S. S. (2005). Measures of working memory, sequence learning, and speech recognition in the elderly. *J. Speech Lang. Hear. Res.* 48, 224–235. doi: 10.1044/1092-4388(2005/016)

Humes, L. E., Kidd, G. R., and Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front. Syst. Neurosci.* 7:55. doi: 10.3389/fnsys.2013.00055

Institute of Electrical, and Electronics Engineers (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust.* 17, 227–246.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 116, 2395–2405. doi: 10.1121/1.1784440

Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J., and Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *Int. J.Otolaryngol.* 2012:865731. doi: 10.1155/2012/865731

Levitt, H., and Resnick, S. B. (1978). Speech reception by the hearing-impaired: methods of testing and the development of new tests. *Scand. Audiol. Suppl.* 6, 107–130.

Luce, P. A. (1986). *Neighborhoods of Words in the Mental Lexicon. Research on Speech Perception*. Technical Report No. 6. Bloomington, IN: Speech Research Laboratory, Department of Psychology, Indiana University.

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42, S49–S58. doi: 10.3109/14992020309074624

Mills, J. H., Schmiedt, R. A., Schulte, B. A., and Dubno, J. R. (2006). Age related hearing loss: a loss of voltage, not hair cells. *Semin. Hear.* 27, 228–236. doi: 10.1055/s-2006-954849

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Parbery-Clark, A., Skoe, E., Lam, C., and Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear. Hear.* 30, 653–661. doi: 10.1097/AUD.0b013e3181b412e9

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., et al. (1996). Mediators of long-term memory performance across the life span. *Psychol. Aging* 11, 621–637. doi: 10.1037/0882-7974.11.4.621

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282

Pisoni, D. B., and Cleary, M. (2004). "Learning, memory, and cognitive processes in deaf children following cochlear implantation," in *Cochlear Implants: Auditory Prostheses and Electric Hearing*, eds F.-G. Zeng, A. N. Popper, and R. R. Fay (New York, NY: Springer), 377–426. doi: 10.1007/978-0-387-22585-2_9

Rabbitt, P. M. (1968). Channel-capacity, intelligibility and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158

Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol.* 111, 167–176. doi: 10.3109/00016489109127274

Rossiter, S., Stevens, C., and Walker, G. (2006). Tinnitus and its effect on working memory and attention. *J. Speech Lang. Hear. Res.* 49, 150–160. doi: 10.1044/1092-4388(2006/012)

Rudner, M., Fransson, P., Ingvar, M., Nyberg, L., and Rönnberg, J. (2007). Neural representation of binding lexical signs and words in the episodic buffer of working memory. *Neuropsychologia* 45, 2258–2276. doi: 10.1016/j.neuropsychologia.2007.02.017

Rudner, M., Rönnberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4

Smith, S. L., Pichora-Fuller, M. K., Wilson, R. H., and MacDonald, E. (2012). Word recognition for temporally and spectrally distorted materials: the effects of age and hearing loss. *Ear. Hear.* 33, 349–366. doi: 10.1097/AUD.0b013e318242571c

Souza, P. E. (2012). *20Q: Cognition Measures– They Might Change the Way You Fit Hearing Aids. Audiology*. Available at: http://www.audiologyonline.com/articles/august-20q-by-pamela-souza-6925.

Tillman, T. W., and Carhart, R. (1966). *An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words* (*Northwestern University Auditory Test No. 6 United States Air Force [USAF] School of Aerospace Medicine Technical Report 66-55*). Brooks Air Force Base, TX: USAF School of Aerospace Medicine.

Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., Hale, S., and Rose, N. S. (2008). Auditory-visual discourse comprehension by older and young adults in favorable and unfavorable conditions. *Int. J. Audiol.* 47, S31–S37. doi: 10.1080/14992020802301662

van Rooij, J. C., and Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II: multivariate analyses. *J. Acoust. Soc. Am.* 88, 2611–2624. doi: 10.1121/1.399981

Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *J. Acoust. Soc. Am.* 107, 1671–1684. doi: 10.1121/1.428451

Waters, G. S., and Caplan, D. (1996).The measurement of verbal working memory capacity and its relation to reading comprehension. *Q. J. Exp. Psychol. A* 49, 51–74. doi: 10.1080/713755607

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (*WAIS–IV*), 4th Edn. San Antonio, TX: NCS Pearson.

Wilson, R. H. (2003). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *J. Am. Acad. Audiol.* 14, 453–470.

Wilson, R. H., Abrams, H. B., and Pillion, A. L. (2003). A word-recognition task in multitalker babble using a descending presentation mode from 24-dB to 0-dB signal to babble. *J. Rehabil. Res. Dev.* 40, 321–328. doi: 10.1682/JRRD.2003.07.0321

Wilson, R. H., McArdle, R., and Roberts, H. (2008). A comparison of recognition performances in speech-spectrum noise by listeners with normal hearing on PB-50, CID W-22, NU-6, W-1 Spondaic Words, and Monosyllabic Digits spoken by the same speaker. *J. Am. Acad. Audiol.* 19, 496–506. doi: 10.3766/jaaa.19.6.5

Wilson, R. H., McArdle, R., Watts, K. L., and Smith, S. L. (2012). The revised speech perception in noise Test (R-SPIN) in a multiple signal-to-noise ratio paradigm. *J. Am. Acad. Audiol.* 23, 590–605. doi: 10.3766/jaaa.23.7.9

Wilson, R. H., Morgan, D. E., and Dirk, D. O. (1973). A proposed SRT procedure and its statistical precedent. *J. Speech Hear. Disord.* 38, 184–191. doi: 10.1044/jshd.3802.184

Wingfield, A. (1996). Cognitive factors in auditory performance: context, speed of processing, and constraints of memory. *J. Am. Acad. Audiol.* 7, 175–182.

Wingfield, A., and Stine-Morrow, E. A. L. (2000). "Language and speech," in *Handbook of Aging and Cognition*, 2nd Edn, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Erlbaum), 359–416.

Zekveld, A. A., Rudner, M., Johnsrude, I. S., and Rönnberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *J. Acoust. Soc. Am.* 134, 2225–2234. doi: 10.1121/1.4817926

Conflict of Interest Statement: The contents of this manuscript do not represent the views of the Department of Veterans Affairs or the United States government. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Smith and Pichora-Fuller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 03 August 2015
doi: 10.3389/fpsyg.2015.01017

How does susceptibility to proactive interference relate to speech recognition in aided and unaided conditions?

Rachel J. Ellis and Jerker Rönnberg*

Department of Behavioural Sciences and Learning, Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

Proactive interference (PI) is the capacity to resist interference to the acquisition of new memories from information stored in the long-term memory. Previous research has shown that PI correlates significantly with the speech-in-noise recognition scores of younger adults with normal hearing. In this study, we report the results of an experiment designed to investigate the extent to which tests of visual PI relate to the speech-in-noise recognition scores of older adults with hearing loss, in aided and unaided conditions. The results suggest that measures of PI correlate significantly with speech-in-noise recognition only in the unaided condition. Furthermore the relation between PI and speech-in-noise recognition differs to that observed in younger listeners without hearing loss. The findings suggest that the relation between PI tests and the speech-in-noise recognition scores of older adults with hearing loss relates to capability of the test to index cognitive flexibility.

Keywords: cognition, speech-in-noise recognition, proactive interference, working memory, executive function, sensorineural hearing loss, hearing aids, older adults

OPEN ACCESS

Edited by:
Anne-Lise Giraud, École Normale Supérieure, France

Reviewed by:
Fatima T. Husain, University of Illinois at Urbana-Champaign, USA
Hyo-Jeong Lee, Hallym University College of Medicine, South Korea

****Correspondence:***
Rachel J. Ellis, Department of Behavioural Sciences and Learning, Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, 58183 Linköping, Sweden
rachel.ellis@liu.se

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *30 January 2015*
Accepted: *06 July 2015*
Published: *03 August 2015*

Citation:
Ellis RJ and Rönnberg J (2015) How does susceptibility to proactive interference relate to speech recognition in aided and unaided conditions? Front. Psychol. 6:1017. doi: 10.3389/fpsyg.2015.01017

Introduction

Proactive interference (PI) refers to an effect whereby the acquisition of new memories is disrupted by interference from similar information that has been learned previously. PI is a robust phenomenon, having been observed in a variety of contexts including memory for odors (Lawless and Engen, 1977) and the probability of developing post-traumatic stress disorder (Verwoerd et al., 2009). However, PI is traditionally investigated in terms of its effects on memory for semantically-related lists of words (see for example: Wickens et al., 1963; Floden et al., 2000; Ellis and Rönnberg, 2014). The earliest studies of PI focussed only on investigating the capacity to resist PI by presenting lists of words to be recalled after a short interval of time. This procedure is known as the Brown–Peterson paradigm (Brown, 1958; Peterson and Peterson, 1959) and has since been modified to also allow for the examination of release from PI. This modified version of the Brown–Peterson task (Wickens et al., 1963; Wickens, 1970) is based on manipulating the semantic categories of the to-be-remembered word lists such that the first three lists belong to the same category (for example, countries) with the final list belonging to a different one (for example, flowers). Using this paradigm, resistance to PI would be operationalised as the difference in performance (that is, number of words correctly recalled) between the three

Abbreviations: HFPTA, High-frequency pure tone average; SIN, Speech in noise; PI, Proactive interference; WM, Working memory.

lists belonging to the same semantic category, with a decrease in performance indicating an effect of PI. The magnitude of release from PI is calculated as the benefit in performance afforded by the change of semantic category between lists three and four.

Effects of PI have been demonstrated in both long-term and short-term memories (Keppel and Underwood, 1962). However, it is the relation between PI and working memory (WM) that is of particular relevance to this study. WM is comprised of both processing and storage components, as opposed to the long-term and short-term memories which simply store information. Thus, rather than simply indexing memory, tests of WM span are thought to measure a number of complex cognitive processes (Sörqvist et al., 2010) including PI (Kane and Engle, 2000; Whitney et al., 2001; Friedman and Miyake, 2004). Studies have also shown that manipulating the degree of PI in tests of WM span affects how well the WM span scores predict performance in other complex cognitive tasks such as tests of prose recall (Lustig et al., 2001) and fluid intelligence (Blalock and McCabe, 2011).

Another complex task, known to be predicted by WM span scores is the perception of distorted speech (see Akeroyd, 2008, for a review). Recent research suggests that performance in a test of PI is significantly related to the speech-in-noise recognition scores of young listeners with normal hearing (Ellis and Rönnberg, 2014). This begs the question of whether the same relation can be observed in older listeners with a hearing loss. According to the ease of language understanding (ELU) model (Rönnberg, 2003; Rönnberg et al., 2008, 2013), when listening conditions are favorable, speech stimuli are implicitly processed, however, if listening conditions are compromised in some way, a mismatch may occur between the stimuli being presented and the representation stored in the long term memory. A mismatch may be caused by many factors, including noise, hearing loss and hearing aid processing and means that explicit processing and storage resources are required, making speech perception more demanding for the listener (Rudner and Rönnberg, 2008). Evidence of increased cognitive load associated with speech perception relative to those with normal hearing has also been observed in cochlear implant users (see for example, Song et al., 2015). Thus, it is expected that a stronger relation between PI and speech-in-noise recognition will be observed in a sample of older listeners with a hearing loss compared to younger listeners with normal hearing. This is due to the fact that the degree of signal distortion, and thus of cognitive resources required to correctly perceive speech-in-noise, is assumed to be greater for older listeners with hearing loss than for younger listeners with normal hearing.

The aim of the present study is therefore to investigate whether the speech-in-noise recognition scores of listeners with an age-related hearing loss is significantly related to performance in a visual PI test. Whether this relation differs depending on whether the speech-in-noise task is completed in an aided or unaided condition will also be investigated, along with the degree to which performance in the PI test relates to aided benefit to speech-in-noise perception.

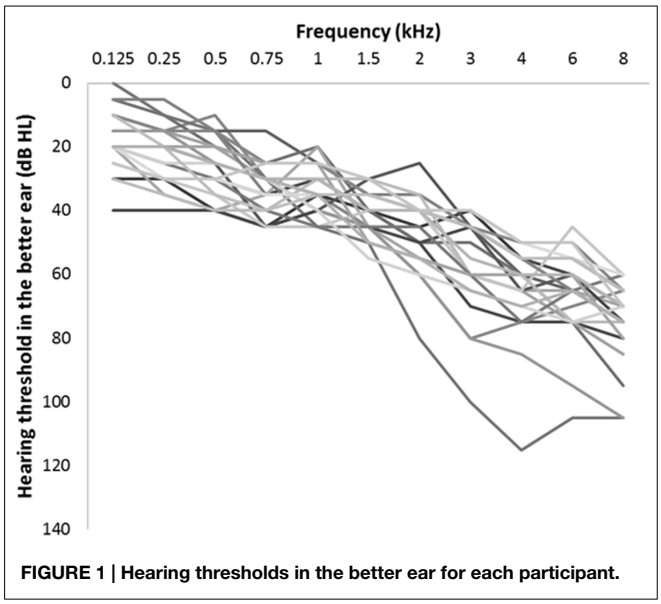


FIGURE 1 | Hearing thresholds in the better ear for each participant.

Materials and Methods

Participants

A sample of 23 participants (16 male) aged between 65 and 77 years old (mean age = 70 years) were recruited via the audiology clinic at Linköping University Hospital to take part in the study. Listeners were required to be native speakers of Swedish and have a moderate—to—severe (in two cases, profound at the high-frequencies) symmetrical sensorineural hearing loss and at least 1 year of hearing aid experience, which was binaural for all participants except one. Participants' better-ear audiograms are displayed in **Figure 1**. Note that, in two cases, thresholds for some of the high-frequency tones exceeded the maximum presentation level of the audiometer. Where this occurred, the maximum presentation level is recorded as the threshold. The study was approved by the Regional Ethics Board in Linköping (Project code: IBL-2013-00208). Participants were paid 500 SEK for taking part in the study.

Procedure

All testing was completed in one session, lasting approximately 1.5 h. Upon arrival, participants completed a questionnaire about their hearing loss and a pure-tone audiogram was obtained (at frequencies between 125 and 8000 Hz). Participants then completed the PI test and finally, the speech-in-noise recognition test. The order in which these tests were completed was not counterbalanced, as it was expected that fatigue could affect performance in either of these tests, thus we wished to keep the order the same for all participants. In order to reduce the potential effects of fatigue, participants were encouraged to take breaks in between the tests.

Speech-in-noise Recognition Test

Six blocks of 10 sentences from the Swedish HINT corpus (Hällgren et al., 2006) were presented at 65 dB SPL via a

loudspeaker situated approximately one meter away from the listener at 0 degrees azimuth. Three of the blocks were presented in an aided condition (using the participant's own hearing aids) and three in an unaided condition. Allocation of each block to the aided or unaided condition was randomized, as was the order in which the conditions were completed. The sentences were presented in a background of 2-talker babble noise at fixed SNRs between +15 and −3 dB increasing in difficulty in 3 dB steps, similar to the method recommended by Wilson et al. (2007). The first three sentences in each block were presented in quiet so as to minimize the threat of floor effects and to help to maintain participants' interest in the task. The participants were asked to listen to one sentence at a time and verbally repeat what they heard back to the experimenter. The outcome measure was the mean percentage of keywords correctly identified. The test took approximately 10–15 min to complete.

Proactive Interference Test

The PI test consisted of three blocks of trials. Each block consisted of four lists of seven words, the first three lists belonging to one semantic category (for example, "capital cities") and the final list belonging to a second semantic category (for example, "birds"). Words were presented orthographically on a computer screen. After the presentation of each list, participants completed a distractor task for 16 s to prevent rehearsal. The distractor task involved participants being presented (orthographically) with a letter-number sequence (for example, "S56") and being asked to continue the sequence ("S57, S58, S59" etc). After the distractor task, participants were given 20 s to recall as many words as possible from the list. Participants gave their answers verbally and their responses were noted down by the experimenter. Two outcome measures were then calculated: Resistance to PI (list 1 recall–list 3 recall), where a lower score indicates greater resistance to PI and Release from PI (list 4 recall–list 3 recall), where a higher score indicates greater release from PI. See **Figure 2** for a depiction of a typical pattern of PI responses.

Prior to analysis, the normality of the data was confirmed, thus parametric tests were conducted. In order to determine whether there was evidence of an effect PI in the data, *t*-tests were used. Correlational analyses were then conducted to investigate the relation between the measures of speech recognition and those of PI. Partial correlations, with the effect of high frequency pure tone average (HFPTA = average hearing threshold across both ears at 2000, 4000, 6000, and 8000 Hz) removed were also conducted to examine the extent to which the relation between the measures of PI and speech recognition was influenced by degree of hearing loss. Reported p-values are based on 1-tailed hypotheses.

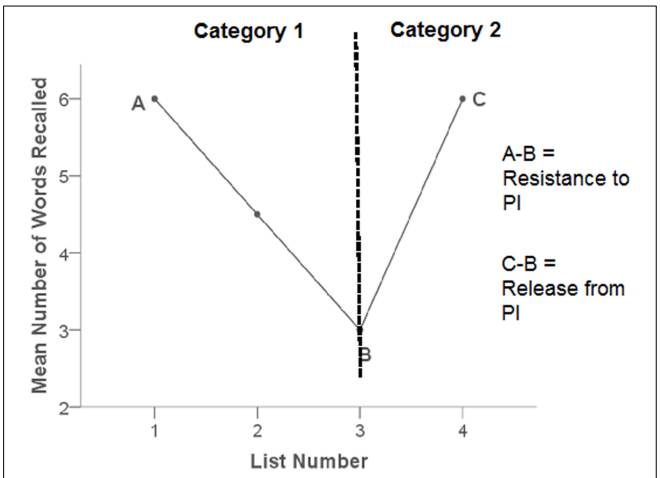


FIGURE 2 | Diagram showing a typical pattern of PI responses.

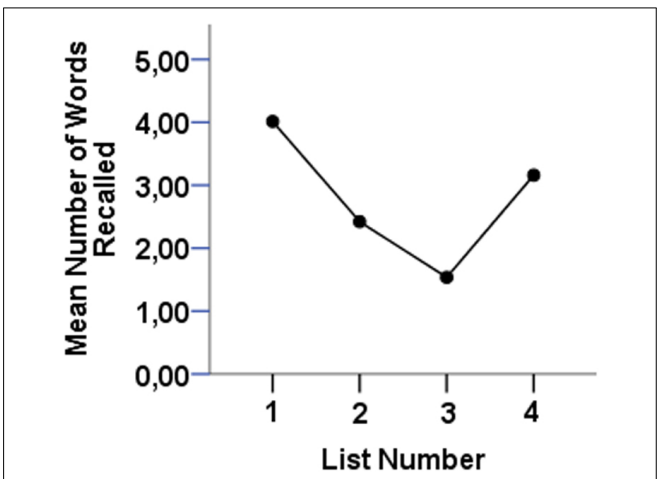


FIGURE 3 | Pattern of recall in the proactive interference test.

Results

Proactive Interference

The mean number of items in each list correctly recalled in the PI task is depicted in **Figure 3**. The results show that performance steadily declines between lists 1 and 3, then increases again at list 4, a pattern consistent with an effect of PI. Paired-samples *t*-tests confirm significant effects of both resistance to PI [$t(68) = 12.34$, $p < 0.000$] and release from PI [$t(68) = 8.42$, $p < 0.000$] thus replicating the expected effects using this task.

Relation Between PI and Speech-in-noise Recognition

SIN Recognition: Unaided

The relation between unaided performance in the SIN test and both resistance to (panel A) and release from PI (panel B) can be seen in **Figure 4**. The results of correlational analyses indicate that only the relation between unaided SIN performance and release from PI is significant ($r = 0.47$, $p = 0.015$), with the relation between unaided SIN performance and resistance to PI failing to reach significance ($r = 0.27$, $p =$ n.s.).

Partial correlational analyses, with the effect of HFPTA removed, revealed the same pattern of results with the relation between unaided SIN and release from PI ($r = 0.46$, $p = 0.015$) showing a significant correlation and the relation between

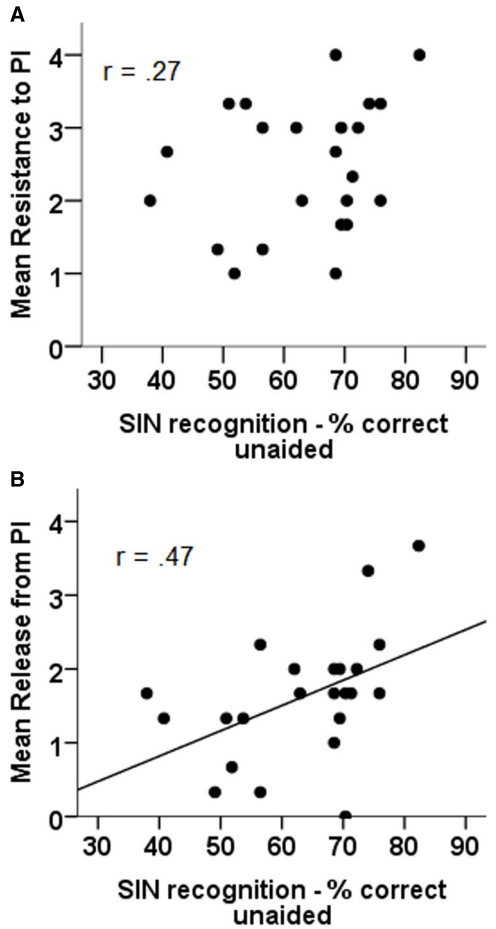


FIGURE 4 | Correlations between unaided sentence in noise recognition and resistance to PI (A) and release from PI (B).

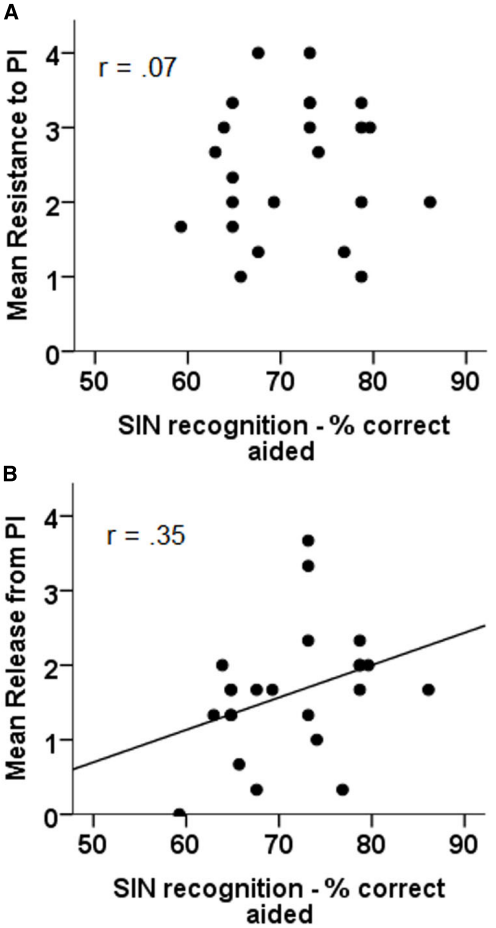


FIGURE 5 | Correlations between aided sentence in noise recognition and resistance to PI (A) and release from PI (B).

unaided SIN and resistance to PI failing to reach significance ($r = 0.26$, $p =$ n.s.).

SIN Recognition: Aided

The relation between aided performance in the SIN test and both resistance to (panel A) and release from PI (panel B) can be seen in **Figure 5**. The results of correlational analyses indicate that only the relation between aided SIN performance and release from PI is significant ($r = 0.35$, $p = 0.05$), with the relation between aided SIN performance and resistance to PI failing to reach significance ($r = 0.07$, $p =$ n.s.).

Once the effect of HFPTA had been removed, the results of the partial correlational analyses indicated that neither the relation between aided SIN performance and release from PI ($r = 0.30$, $p =$ n.s.) nor the relation between aided SIN performance and resistance to PI ($r = -0.19$, $p =$ n.s.) were significant.

SIN recognition: Aided benefit

The relation between aided benefit in the SIN test and both resistance to (panel A) and release from PI (panel B) can be seen in **Figure 6**. The results of correlational analyses indicate that neither the relation between aided benefit in the SIN test and release from PI ($r = -0.25$, $p =$ n.s.) nor the relation between aided benefit in the SIN test and resistance to PI ($r = -0.22$, $p =$ n.s.) were significant.

Partial correlational analyses, with the effect of HFPTA removed, revealed the same pattern of results with neither the relation between aided benefit in the SIN test and release from PI ($r = -0.25$, $p =$ n.s.) nor the relation between aided benefit in the SIN test and resistance to PI ($r = -0.22$, $p =$ n.s.) reaching significance.

Discussion

The results of the study provide clear evidence of resistance to and release from PI on a semantically-based word recall task, based on the modified Brown–Peterson paradigm (Brown, 1958; Peterson and Peterson, 1959). The findings indicate that release from PI is significantly correlated with both aided and unaided speech-in-noise recognition in older listeners with hearing loss. Furthermore, the relation between PI and unaided speech recognition continues to be significant even when the effects of loss of high-frequency hearing sensitivity are removed. However, performance on the PI task did not correlate significantly with the

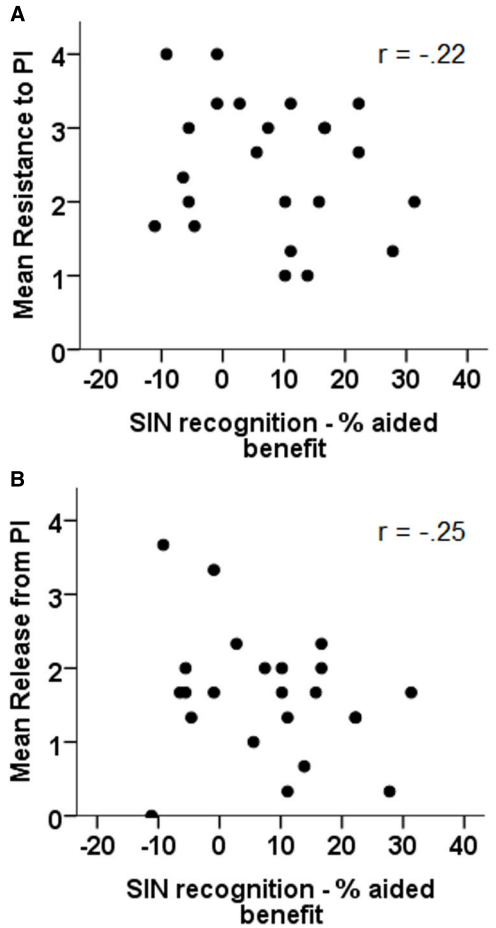


FIGURE 6 | Correlations between aided benefit in the sentence in noise recognition test and resistance to PI (A) and release from PI (B).

degree of benefit to speech-in-noise recognition provided by the use of hearing aids.

Evidence of Proactive Interference

The results of the study provide evidence of significant effects of both resistance to and release from PI. The magnitude of this effect was greater than that observed in our earlier study on younger listeners with normal hearing (Ellis and Rönnberg, 2014). This difference is likely due to the difference in age of the participant groups with older participants being more affected by interference than younger listeners (see for example, Pettigrew and Martin, 2014). In addition, it is plausible that the nature of the distractor task may have put the older participants at a disadvantage compared to the younger participants as older participants have more difficulty completing tasks involving task switching (see for example, Lawo et al., 2012).

It may also be that, despite the fact that the test of PI used in this case contained no auditory information, that listeners with a hearing loss were disadvantaged anyway, due to the association between hearing loss and cognitive decline (see for example,Rönnberg et al., 2011, 2014). However, as we did not include a control group of older listeners without hearing loss it is difficult to determine whether this is in fact the case.

Due to differences in the methodologies employed, it is difficult to draw direct comparisons between the magnitude of the effects of PI observed in the present study and the results reported in previous studies. However, the only methodological difference between this study and that reported by Ellis and Rönnberg (2014) is that stimuli were presented orthographically rather than aurally as was the case in the earlier study. Thus, it may be that had listeners in our previous study been given the orthographic version of the test, they would have been affected by PI to a greater degree than that observed.

Proactive Interference and Speech in Noise Recognition

The results of the study indicate that, in the case of older listeners with hearing loss, release from PI correlates significantly with both aided and unaided speech in noise. This pattern of results differs to that observed in younger adults without hearing loss, for whom resistance to, rather than release from, PI was significantly related to speech-in-noise recognition. Furthermore, the magnitude of the observed effects of both resistance to, and release from, PI were greater in the present study than in our earlier study on young adults with normal hearing (Ellis and Rönnberg, 2014).

One possible explanation for these results may relate to the fact that older adults are known to have a greater bias to respond in a context-congruent manner and be less able than younger adults to constrain responses to a given category (Rogers et al., 2012). These tendencies may contribute both to the larger PI effects observed in this older group, and to the difference in how the effects of PI relate to speech-in-noise recognition. We suggest that in both younger and older adults, resistance to PI provides a measure of the capacity to inhibit interference or to direct attention to specific stimuli, capacities which are sufficient to correlate significantly with how well a younger person is able to recognize speech in noise. However, in the case of older adults with hearing loss, we hypothesize that this capacity alone is not sufficient to predict speech in noise recognition, due to fact that speech recognition is more cognitively taxing for this group than for younger adults. Thus we suggest that, in older adults, release from PI may provide an index of the ability to deviate from context, in essence a measure of cognitive flexibility. If this is the case, we would expect release from PI to correlate more strongly with speech-in-noise recognition in conditions in which less contextual information is available, and indeed our results suggest that this is the case. Specifically, once the effects of loss of high frequency hearing sensitivity had been partialled out, release from PI continued to correlate significantly with unaided speech perception, however, ceased to correlate significantly with aided speech recognition. It should be noted that we have made no attempt to disentangle the effects of aging and hearing loss in our data, thus our findings reflect the combined influence of both factors. However, recent research suggests that even when older listeners have normal audiometric thresholds, they tend to perform more poorly on speech perception tests than do younger participants (Füllgrabe et al., 2014).

That being the case, we would hypothesize that PI is likely to correlate with speech in noise perception in older adults without hearing loss, however, further research is necessary to investigate this issue.

The fact that that release from PI correlates with speech in noise perception in the unaided condition only is consistent with the ELU model (Rönnberg, 2003; Rönnberg et al., 2008, 2013) if we assume that hearing aids decrease distortion of the signal and allow for more implicit, relatively cognitively undemanding, processing of speech as opposed to the explicit, more cognitively demanding, processing of unaided speech which may be perceived as distorted and inconsistent with representations stored in the long term memory. However, neither measure of PI correlated significantly with the degree of benefit to speech recognition afforded by hearing aid use. There are a number of methodological reasons that may explain this seemingly inconsistent finding. The first is that we did not check how well the hearing aids matched the participants' prescription. Furthermore, we were unable to check which signal processing options were active in the participants' hearing aids. There are a number of studies that have linked cognitive status to the success or lack thereof of a particular signal processing strategy to an individual listener (Lunner, 2003; Rudner et al., 2008). Thus it may be that, taken together, these methodological issues may have obscured the relation between PI and aided benefit to speech perception. It may also be of interest to investigate whether the relation between PI and unaided speech perception is affected by regular use of hearing aids, which may affect the degree to which the unaided representations (mis)match with the representations stored in the long-term memory.

The results seem to indicate both that PI is involved in speech perception and that hearing aids facilitate a decreased reliance on cognitive function. The findings seem to be inconsistent with the suggestion that release from PI is an automatic process and unrelated to WM capacity (see Kane and Engle, 2000; Friedman and Miyake, 2004). In the present study, we observe that resistance to and release from PI are significantly correlated with each other indicating that release from PI, at least as measured in the present study, does not simply reflect an automatic process but rather a more explicit process as is the case with resistance to PI. Furthermore, the fact that, after correction for HFPTA, release from PI correlates with speech perception in only the unaided condition, gives further support to the idea that release from PI may be a more complex process that previously thought.

Author Contributions

RE and JR contributed equally to designing the study, interpreting the data and preparing the manuscript. RE collected the data.

Acknowledgments

This research was funded by the Linnaeus Centre HEAD, The Swedish Research Council (grant number: 2007-8654). The authors would like to thank Amin Saremi for technical assistance and Mathias Hällgren for help recruiting participants. The authors would also like to thank the participants for donating their time. Parts of this work were presented at the International Hearing Aid Research Conference, Lake Tahoe, USA, August 13–17 2014.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Blalock, L. D., and McCabe, D. P. (2011). Proactive interference and practice effects in visuospatial working memory span task performance. *Memory* 19, 83–91. doi: 10.1080/09658211.2010.537035

Brown, J. A. (1958). Some tests of the decay theory of immediate memory. *Q. J. Exp. Psychol.* 10, 12–21.

Ellis, R. J., and Rönnberg, J. (2014). Cognition and speech-in-noise recognition: the role of proactive interference. *J. Am. Acad. Audiol.* 25, 975–982. doi: 10.3766/jaaa.25.10.6

Floden, D., Stuss, D. T., and Craik, F. I. M. (2000). Age differences in performance on two versions of the Brown–Peterson task. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 7, 245–259. doi: 10.1076/anec.7.4.245.795

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2014). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347. doi: 10.3389/fnagi.2014.00347

Hällgren, M., Larby, B., and Arlinger, S. (2006). A Swedish version of the hearing in noise test (HINT) for measurement of speech recognition. *Int. J. Audiol.* 45, 227–237. doi: 10.1080/14992020500429583

Kane, M. J., and Engle, R. W. (2000). Working-memory capacity, proactive interference and divided attention: limits on long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 336–358. doi: 10.1037/0278-7393.26.2.336

Keppel, G., and Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *J. Verbal Learn. Verbal Behav.* 1, 153–161. doi: 10.1016/S0022-5371(62)80023-1

Lawless, H., and Engen, T. (1977). Associations to odors: interference, mnemonics, and verbal labeling. *J. Exp. Psychol. Hum. Learn.* 3, 52–59. doi: 10.1037/0278-7393.3.1.52

Lawo, V., Philipp, A. M., Schuch, S., and Koch, I. (2012). The role of task preparation and task inhibition in age-related task-switching deficits. *Psychol. Aging* 27, 1130–1137. doi: 10.1037/a0027455

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42, S49–S58. doi: 10.3109/14992020309074624

Lustig, C., May, C. P., and Hasher, L. (2001). Working memory span and the role of proactive interference. *J. Exp. Psychol. Gen.* 130, 199–207. doi: 10.1037/0096-3445.130.2.199

Peterson, L. R., and Peterson, M. J. (1959). Short-term retention of individual items. *J. Exp. Psychol.* 58, 193–198. doi: 10.1037/h0049234

Pettigrew, C., and Martin, R. C. (2014). Cognitive declines in healthy aging: evidence from multiple aspects of interference resolution. *Psychol. Aging* 29, 187–204. doi: 10.1037/a0036085

Rogers, C. S., Jacoby, L. L., and Sommers, M. S. (2012). Frequent false hearing by older adults: the role of age differences in metacognition. *Psychol. Aging* 27, 33–45. doi: 10.1037/a0026231

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42, S68–S76. doi: 10.3109/14992020309074626

Rönnberg, J., Danielsson, H., Rudner, M., Arlinger, S., Sternang, O., Wahlin, A., et al. (2011). Hearing loss is negatively related to episodic and semantic long-term memory but not to short-term memory. *J. Speech Lang. Hear Res.* 54, 705–726. doi: 10.1044/1092-4388(2010/09-0088)

Rönnberg, J., Hygge, S., Keidser, G., and Rudner, M. (2014). The effect of functional hearing loss and age on long-and short-term visuospatial memory:

evidence from the UK biobank resource. *Front. Aging Neurosci.* 6:326. doi: 10.3389/fnagi.2014.00326

Rönnberg, J., Lunner, L., Zekveld, A., Sörqvist, P., Danielsson, H., Signoret, C., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system of ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi: 10.1080/14992020802301167

Rudner, M., Foo, C., Sundewall-Thorén, E., Lunner, T., and Rönnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing aid users. *Int. J. Audiol.* 47, S91–S98. doi: 10.1080/14992020802304393

Rudner, M., and Rönnberg, J. (2008). The role of the episodic buffer in working memory for language processing. *Cogn. Process.* 9, 19–28. doi: 10.1007/s10339-007-0183-x

Song, J.-J., Lee, H.-J., Kang, H., Lee, D. S., Chang, S. O., and Oh, S. H. (2015). Effects of congruent and incongruent visual cues on speech perception and brain activity in cochlear implant users. *Brain Struct. Funct.* 220, 1109–1125. doi: 10.1007/s00429-013-0704-6

Sörqvist, P., Ljungberg, J. K., and Ljung, R. (2010). A subprocess view of working memory capacity: evidence from effects of speech on prose memory. *Memory* 18, 310–326. doi: 10.1080/09658211003601530

Verwoerd, J., Wessel, I., and de Jong, P. J. (2009). Individual differences in experiencing intrusive memories: the role of the ability to resist proactive interference. *J. Behav. Ther. Exp. Psychiatry* 40, 189–201. doi: 10.1016/j.jbtep.2008.08.002

Whitney, P., Arnett, P. A., Driver, A., and Budd, D. (2001). Measuring central executive functioning: what's in a reading span? *Brain Cogn.* 45, 1–14. doi: 10.1006/brcg.2000.1243

Wickens, D. D. (1970). Encoding categories of words: an empirical approach to meaning. *Psychol. Rev.* 77, 1–15. doi: 10.1037/h0028569

Wickens, D. D., Born, D. G., and Allen, C. K. (1963). Proactive inhibition and item similarity in short term memory. *J. Verbal Learn. Verbal Behav.* 2, 440–445. doi: 10.1016/S0022-5371(63)80045-6

Wilson, R. H., McArdle, R. A., and Smith, S. L. (2007). An evaluation of the BKB-SIN, HINT, QuickSIN and WIN materials on listeners with normal hearing and listeners with hearing loss. *J. Speech Lang. Hear. Res.* 50, 844–856. doi: 10.1044/1092-4388(2007/059)

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Ellis and Rönnberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 22 July 2015
doi: 10.3389/fpsyg.2015.00987

Keeping track of who said what: Performance on a modified auditory *n*-back task with young and older adults

Gary R. Kidd and Larry E. Humes*

Department of Speech and Hearing Sciences, Indiana University, Bloomington, IN, USA

A modified auditory *n*-back task was used to examine the ability of young and older listeners to remember the content of spoken messages presented from different locations. The messages were sentences from the Coordinative Response Measure (CRM) corpus, and the task was to judge whether a target word on the current trial was the same as in the most recent presentation from the same location (left, center, or right). The number of trials between comparison items (the number back) was varied while keeping the number of items to be held in memory (the number of locations) constant. Three levels of stimulus uncertainty were evaluated. Low- and high-uncertainty conditions were created by holding the talker (voice) and nontarget words constant, or varying them unpredictably across trials. In a medium-uncertainty condition, each location was associated with a specific talker, thus increasing predictability and ecological validity. Older listeners performed slightly worse than younger listeners, but there was no significant difference in response times (RT) for the two groups. An effect of the number back (*n*) was seen for both PC and RT; PC decreased steadily with *n*, while RT was fairly constant after a significant increase from $n = 1$ to $n = 2$. Apart from the lower PC for the older group, there was no effect involving age for either PC or RT. There was an effect of target word location (faster RTs with a late-occurring target) and an effect of uncertainty (faster RTs with a constant talker-location mapping, relative to the high-uncertainty condition). A similar pattern of performance was observed with a group of elderly hearing-impaired listeners (with and without shaping to ensure audibility), but RTs were substantially slower and the effect of uncertainty was absent. Apart from the observed overall slowing of RTs, these results provide little evidence for an effect of age-related changes in cognitive abilities on this task.

Keywords: hearing, speech perception, effort, working memory, aging

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Bruce A. Schneider,
University of Toronto Mississauga, Canada
Sherri L. Smith,
Veterans Affairs Medical Center, USA

****Correspondence:***
Gary R. Kidd,
Department of Speech and Hearing Sciences, Indiana University, 200 S. Jordan Ave., Bloomington, IN 47405, USA
kidd@indiana.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *04 March 2015*
Accepted: *29 June 2015*
Published: *22 July 2015*

Citation:
Kidd GR and Humes LE (2015) Keeping track of who said what: Performance on a modified auditory n-back task with young and older adults. Front. Psychol. 6:987. doi: 10.3389/fpsyg.2015.00987

Introduction

Many people have difficulty participating in conversations when listening conditions are not ideal. Speaking with one person face-to-face in a quiet environment is considerably easier than conversing with a group in a noisy restaurant, and the difficulty tends to be greater for older listeners, especially those with hearing loss (Humes, 2002; Humes and Dubno, 2010). Many factors can make listening

conditions more difficult, including background noise, competing speech sounds, reverberation, poor enunciation, and other types of distraction or signal degradation (see Mattys et al., 2012, for a recent review). One very common factor, often confounded with background noise, is the presence of multiple talkers participating in a group conversation. Such conversations often take place in noisy environments with the participants talking over each other; but even in a quiet environment with polite turn taking, this can be a challenging listening situation for some listeners. Following a sequential or turn-taking multitalker conversation generally requires a listener to keep track of interleaved remarks from multiple talkers. Listening to each new contribution to a conversation while remembering earlier remarks from other talkers and correctly attributing those remarks to the different participants places demands on cognitive abilities that tend to decline with age.

Many studies have shown that cognitive abilities generally diminish with age (see Salthouse, 2010, 2012, for reviews), but the degree to which this reduces the ability to follow a multitalker conversation is not clear. The task of following a conversation consisting of interleaved messages from multiple talkers may involve several distinct abilities. Although a multitalker conversation puts demands on working memory, in that listeners are required to keep information in memory while processing new information from each new talker, it also involves other abilities that may be independent of the simultaneous memory and processing abilities assessed by working memory tasks. For example, localization ability, selective listening abilities, the ability to make use of partial information, and the ability to deal with uncertainty may all come into play in a multitalker listening situation. These abilities may be largely independent and may not be affected by aging in the same way.

The implications of age-related changes in cognitive abilities for speech understanding are not always clear. When present, hearing loss is often the primary reason for a decrease in speech-understanding ability with increasing age, but cognitive factors also play an important role. The role of cognitive factors is most apparent when speech is presented in a background of competing speech or speech-like sounds and when the speech is amplified to ensure audibility (see Akeroyd, 2008; Houtgast and Festen, 2008; Humes and Dubno, 2010; Humes et al., 2013). There are very large individual differences in speech-understanding abilities at all ages, so one must be careful about generalizations concerning the abilities of younger vs. older listeners. Even with audible (amplified) speech, older listeners often perform worse than younger listeners under difficult listening conditions (e.g., Humes et al., 2006; Humes and Coughlin, 2009; Kidd and Humes, 2012). However, with fully audible speech, older subjects also perform as well as younger listeners on many difficult speech-understanding tasks (e.g., Humes et al., 2013). Moreover, with highly predictable speech stimuli that provide linguistic and prosodic context, older listeners often outperform younger listeners (e.g., Pichora-Fuller et al., 1995; Wingfield et al., 2000; Humes et al., 2013).

Given the large individual differences in speech-understanding abilities and the dependency of age effects on the specifics of the speech task, it is difficult to predict how age-related changes in hearing and cognition will affect performance in more complex everyday listening situations. Much of what is known about the influence of hearing loss and cognitive abilities on speech understanding comes from studies that require subjects to recall words immediately after presentation of a single word or sentence. However, in everyday listening situations, successful communication requires more than recognition and immediate recall. Although some researchers have examined age differences in the performance of more complex speech-understanding tasks (e.g., Pichora-Fuller et al., 1995; Schurman et al., 2014), much remains unknown about how older listeners are affected by the increased cognitive demands of real-world conversational tasks.

The present study uses an approximation of a multitalker sequential conversation to examine the influence of several factors on the ability to understand and recall information in a series of spoken sentences. To assess the role of aging and hearing loss on this task, the study employs young, normal-hearing (YNH) adults, and older adults, with and without hearing loss. A modified auditory *n*-back task was used with spoken sentences as stimuli. This paradigm, described in more detail below, provides a means of assessing memory for words in different sentence positions under different levels and types of uncertainty, or variability, across trials. The *n*-back task provides a convenient framework for examining these variables in an experimental paradigm that has many features in common with a sequential or turn-taking multitalker conversation.

The *n*-back Task

The *n*-back task is widely used as a measure of working memory, especially in cognitive neuroscience research (e.g., Cohen et al., 1997; Owen et al., 2005). The task requires subjects to judge whether information presented on the current trial matches that presented on an earlier trial, one or more (n) trials back in a sequence of trials. To perform this task, a subject must hold the last n items in memory, so that the identity of the item n presentations prior to the current one is always available as new items are presented. For this basic version of the task, n is therefore equal to both the number of presentations back in the sequence for the comparison item and the memory set size. The task is typically performed in the visual modality with single letters or digits presented individually in a sequence. Many variants of this task (including different presentation strategies, stimuli, and presentation modalities) have been used to test various hypotheses concerning control processes and memory systems in working memory (e.g., McElree, 2001; Oberauer and Bialkova, 2009; Basak and Verhaeghen, 2011; see Owen et al., 2005; Redick and Lindsey, 2013, for reviews). Like any working memory task, the *n*-back task has some task-specific demands that involve abilities that may have little or nothing to do with the basic processing and capacity aspects of working memory (see Kane et al., 2007; Schmiedek et al., 2009). Moreover, *n*-back tasks have been shown to have a fairly weak correlation with other measures of working memory that consist of interleaved memory and processing tasks (Redick and Lindsey, 2013). These "complex-span" tasks (e.g., reading span, operation span; see Conway et al., 2005, for a summary) have been more popular than *n*-back tasks as measures of working memory in most

research on individual differences in cognition (e.g., Daneman and Merikle, 1996; Unsworth and Engle, 2007). The substantial differences in performance on various working memory tasks show that working memory is a complex construct that cannot be effectively assessed with a single measure. However, the *n*-back task has some properties that make it useful for the assessment of certain features of working memory in the context of a sequential multitalker conversation.

In the present study, a modified *n*-back task is used to measure the ability to recall information in sentences spoken by different talkers at different times in a series of spoken messages. The use of this type of task makes it possible to assess components of working memory (such as focus switching and memory for items outside the focus of attention) and determine their influence on the ability to follow a multitalker conversation. The modification of the *n*-back task used here is similar to that used by Verhaeghen and Basak (e.g., Verhaeghen and Basak, 2005; Basak and Verhaeghen, 2011) and Oberauer (2002, 2006) in their investigations of working-memory processes with a visual *n*-back task. Their work has examined effects of aging on the ability to switch items stored in memory in and out of the focus of attention (focus switching) and the probability of recalling items stored outside of the focus of attention (item availability). The research is guided by a two-stage model of working memory (see Cowan, 2001) that posits two memory stores in working memory: a very limited capacity store that affords immediate access (i.e., the focus of attention) and a larger "outer" store in which items are in an activated or available state, but not accessible until they are brought into the focus of attention (with a "focus switch"). With the *n*-back task, focus switching and item search time (for items in the outer store) can be assessed by measuring the time required to judge whether the current item was repeated *n* trials ago as a function of *n* (which is also equal to the number of items that must be held in memory to perform the task). Assuming a 1-item capacity for focus of attention, response times (RT) for $n = 1$ trials can be compared to that for $n = 2$ trials to obtain a measure of switching time. This is because on one-back trials, the current item is compared to the immediately preceding item, an item that is still held in the focus of attention, while on two-back trials, an item must be switched from the outer store into the focus of attention. Any increase in RT with further increases in *n* indicates search time for items in the outer store. In addition to memory search efficiency, the *availability* of items in the outer store can be assessed by examining percent-correct performance as a function of *n*.

The standard version of the *n*-back task has two characteristics that make it unsuitable for assessing sequential multitalker-conversation abilities. First, in conversation we need to keep track of who said what, but a precise ordering of the different participants' contributions to the conversation is generally not important, as long as we follow the flow of the conversation. That is, we can generally follow a conversation quite well even if we are not sure whether two or three other people have spoken since the person sitting next to us last spoke. Second, people do not contribute to a conversation in a fixed order, with everyone contributing once before contributing again. Both of these constraints can be eliminated with an auditory *n*-back task simply by presenting stimuli from fixed locations and asking subjects to judge whether the stimulus they just heard in a given location is the same as the last one they heard in that same location. With this task, the number of items to be remembered (or set size) is equal to the number of locations used. Further, if stimuli are presented from different locations in an unpredictable order, the number of trials between the current stimulus and the comparison stimulus (i.e., the number back, *n*) can be varied independently of the set size and can even exceed the set size.

This type of auditory *n*-back task is illustrated in **Table 1** using a set size of 3 (i.e., three locations). As illustrated, the subject must retain both what was said (a spoken digit in this example) and from where it originated (left, center, or right). With three locations, the subject must remember only three digits and simply indicate "yes" or "no" to indicate whether the digit just heard matches the last digit heard from the same location. As noted, the accuracy of the responses is recorded together with the RTs and both are examined as a function of *n*.

Verhaeghen and Basak (Verhaeghen and Basak, 2005; Basak and Verhaeghen, 2011) have used a visual *n*-back task that shares some properties with the present task. For a set size of one, when focus switching is never required, older adults were as accurate as younger adults and performed nearly perfectly. However, for set sizes greater than one, older adults were less accurate than younger ones on one-back trials (not requiring a focus switch) as well as on trials that did require a focus switch (i.e., trials with a comparison more than one back in the series of presentations). Thus, the burden of keeping track of more than one location for target numbers (and/or maintaining one or more items in the outer store) had a negative effect on older adults' performance, even on trials that did not require focus switching. This shows that, at least under some conditions, older adults have more difficulty maintaining information both inside and outside the focus of attention than do younger adults. However, no differences were found between young and older subjects in focus-switching costs, measured by response times,

TABLE 1 | An example of 10 trials of an auditory *n*-back task with spoken digits presented from three locations; left, center, or right.

Trial	LEFT	CENTER	RIGHT	Trial type
1	"two"			No
2		"six"		No
3			"seven"	No
4		"six"		$n = 2$, Yes
5	"three"			$n = 4$, No
6	"three"			$n = 1$, Yes
7		"five"		$n = 3$, No
8		"eight"		$n = 1$, No
9			"seven"	$n = 6$, Yes
10	"four"			$n = 4$, No

The subjects task is simply to indicate (yes or no) whether the digit just heard is the same as the previous digit heard from that same location. The value of n is the number of trials back in the sequence for the comparison.

when general slowing was taken into account. That is, the relative increase in RTs between one-back trials and two-back (or greater) trials was approximately the same for younger and older subjects.

The present studies provide measures of these working-memory processes (i.e., focus switching, memory search, and availability of information outside focal attention) in the context of an auditory *n*-back task that has some of the properties of a sequential multitalker conversation. Similar to the illustration in **Table 1**, full sentences are presented auditorily from different apparent locations (left, center, or right) over headphones. Subjects are asked to judge whether a target word in the sentence they just heard is the same as that in the most recent sentence presented from the same location. This creates a more natural task that resembles the task of listening to three people (separated in space) and keeping track of who said what.

Although subjects are asked to remember (and compare) only one key word in each sentence, the additional information in the full sentence adds to the processing burden and is potentially distracting. Moreover, the use of apparent location to indicate the stimulus to be compared to the current stimulus may not be as effective as column position in a visual display, as used by Verhaeghen and Basak (2005), because of both age-related changes in localization ability (see Dobreva et al., 2011) and differences between memory for auditorily specified location and memory for location in a two-dimensional visual array (see Parmentier and Jones, 2000; Martin et al., 2011).

In addition to the changes in modality and stimulus complexity, the current studies also differ from earlier *n*-back studies by including a manipulation of the variability in the sentences across trials as a way to measure effects of complexity and stimulus uncertainty on performance. This includes a condition in which the same sentence spoken by the same talker is used across trials with only a change in the key word (minimum uncertainty), plus a condition with variation in talkers and sentences across trials (maximum uncertainty). A third condition more closely approximates a real sequential multitalker situation by having a constant talker-location correspondence while maintaining the same stimulus variability as the maximum uncertainty condition. This medium-uncertainty condition provides a test of the potential benefit due to the ecological validity of each location being associated with a different specific voice (or person) as well as the potential benefit due to comparisons of words spoken in the same voice.

Although the modified *n*-back task used in this study does not have all of the characteristics of a real sequential multitalker conversation, the task and the various conditions used allow for tests of the role of many factors that may play an important role in the ability to follow a real-world multitalker conversation. These include focus switching, memory search, the availability of items in memory (outer store), cognitive load, distraction, uncertainty, the use of location cues, and the use of indexical properties of speech. To determine how these factors are affected by hearing loss and aging, two experiments were conducted: one with young and older adults with normal hearing, and one with older hearing-impaired adults tested with and without spectral shaping (amplification) to ensure full audibility of the stimuli.

Experiment 1: Young and Older Adults with Normal Hearing

The first experiment examined performance on a modified auditory *n*-back task by younger and older adults with normal hearing. Based on performance with a similar visual task (see Oberauer, 2006; Basak and Verhaeghen, 2011), it was expected that older subjects would be slower and less accurate than younger subjects, but that the two groups would have similar switching costs, as evidenced by the relative increase in RT from $n = 1$ (when no focus switching is required) to $n = 2$ (when focus switching is required). The increased processing load due to the use of full sentences, rather than single letters or numerals, was expected to have a greater impact on the older listeners. This would lead to larger age differences in percent-correct performance than seen in related earlier studies with simpler stimuli, and possibly to reduced efficiency in memory search, which would tend to increase RTs on trials with $n > 1$, due to slower searching for items in the outer store. The use of target words early and late in the sentence provides a test of potential memory interference due to irrelevant information preceding or following the target word. Finally, the use of the different uncertainty conditions provides a test of the effect of stimulus variability on younger and older listeners as well as a test of the possible benefit due to the consistent mapping of voices to locations, which more closely approximates an everyday listening situation.

Methods

Subjects

Two groups of listeners participated in Experiment 1. The young, normal-hearing (YNH) group consisted of 10 young adults (3 men and 7 women) between the ages of 20 and 24 years (mean = 22.2 years; $SD = 1.3$). The older normal-hearing (ONH) group consisted of 12 older adults (6 men and 6 women) between the ages of 61 and 72 years (mean = 66.2 years; $SD = 3.5$). All YNH listeners had pure tone thresholds $\leq$ 25dB HL (ANSI, 2004) for all octave frequencies between 250 and 8000 Hz. ONH listeners were required to have a pure tone average ($PTA_{500,1000,2000\,Hz}$) $\leq$ 15 dB HL and a high-frequency PTA ($HFPTA_{1000,2000,4000\,Hz}$) $\leq$ 25 dB HL. All subjects had normal tympanograms and otoscopic findings and showed no evidence of middle ear pathology. The YNH subjects were students at Indiana University in Bloomington and the ONH subjects were from the Bloomington, Indiana community. The ONH subjects (with 2 exceptions) had served in an earlier individual differences study (Humes et al., 2013), which had included screening for serious cognitive and physical impairment. The highest level of education completed ranged from high school (one subject) to vocational school (two subjects), college (five subjects), and graduate school (four subjects). All subjects were native speakers of English and were paid for their participation. Subject recruitment and all experimental procedures were reviewed and approved by the IRB at Indiana University.

Stimuli

The stimuli were sentences from the Coordinate Response Measure (CRM) Corpus (Bolia et al., 2000). This corpus consists of a collection of sentences spoken by four male and four female talkers. All sentences are of the form "Ready [call sign] go to [color] [number] now." There are eight call signs (arrow, baron, charlie, eagle, hopper, laker, ringo, tiger), four colors (blue, green, red, white), and eight numbers (1–8) spoken in all 256 combinations by each talker. Three talkers (two male and one female), judged to be maximally distinguishable by three research assistants, were selected for this study.

Stimuli were presented at 85 dB SPL. This relatively high presentation level was used to approximate the levels used with the older hearing-impaired (OHI) listeners in Experiment 2. For those listeners, the stimuli were amplified to ensure audibility (at least 13 dB above threshold) for frequencies from 125 to 4000 Hz, often resulting in presentation levels above 80 dB SPL. Previous work has shown that presentation levels in this range generally lead to slightly poorer intelligibility (e.g., Dubno et al., 2005a,b; Studebaker et al., 1999) in normal-hearing listeners.

Procedures

All testing was done in a sound-treated booth that met or exceeded ANSI guidelines for permissible ambient noise for earphone testing (American National Standards Institute, 1999). Stimuli were presented binaurally, using Etymotic Research ER-3A insert earphones. Stimuli were presented by computer using a Digital Audio Labs Card Deluxe sound card and a Tucker Davis Technologies System-3 HB7 headphone buffer. Each listener was seated in front of a touchscreen monitor, with a keyboard and mouse available.

On each trial, a single CRM sentence was presented to the left, right, or both earphones to simulate left, right, or center locations, respectively, for the apparent source. All subjects reported that the three apparent source locations were easily identified. Each trial began with the word "LISTEN" presented visually on the display, followed 500 ms later by presentation of a sentence. After each presentation, subjects responded by touching (or clicking with a mouse) one of two virtual buttons (labeled "yes" and "no") on a touch screen display to indicate whether the target word (either the number or the call sign) was the same as that spoken by the last talker heard from the same location. The next trial was presented immediately after the subject responded. No feedback was provided (except during practice trials, described below). Subjects were told to respond as quickly as possible without making errors and were encouraged to guess when they felt unsure of the correct response.

A trial block consisted of a sequence of 33 trials with location repetitions beginning on the fourth trial. An example of the first 10 trials of a block is shown in **Table 2**, with number as the target word. The first three trials were always presented in the left, center, and right virtual locations, in that order, and subjects were instructed to respond "no" to those trials (the "yes" option did not appear) since there was no repetition of any location. This resulted in 30 observations per trial block. The contents of each of the 33 trials in a block were randomized with the following constraints. Within the sequence of 33 trials, each virtual location was used 11 times. The number of trials since the last presentation in a given location (n) ranged from 1 to 5, with 6 repetitions of each value of n in each block of trials. Each of the 8 target words (call signs or numbers) was used at least twice and no more than 8 times within a trial block. All subjects began with four practice trial blocks: two with the number target, followed by two with the call sign target. During the practice trials, correct/incorrect feedback was provided on every trial.

In three different uncertainty conditions, the selection of non-target words in the sentences and the assignment of talkers (voices) to different locations were varied. (See **Table 2** for an example of nontarget word variation.) In the low uncertainty condition, the same voice was used on every trial (the same male voice for all subjects) and all words in the sentence other than the target word (call sign or number) were the same on every trial. In the high uncertainty condition, the talker and the two variable nontarget words (color and either call sign or number) were selected randomly on each trial. A third, more ecologically valid, condition had the same random variation in nontarget

TABLE 2 | An example of 10 trials of the modified auditory *n*-back task, with number as the target word.

Trial	LEFT	CENTER	RIGHT	Trial Type
1	Ready Baron go to blue TWO now			No
2		Ready Ringo go to red SIX now		No
3			Ready Hopper go to white SEVEN now	No
4		Ready Baron go to blue SIX now		$n = 2$, Yes
5	Ready Eagle go to green THREE now			$n = 4$, No
6	Ready Charlie go to white THREE now			$n = 1$, Yes
7		Ready Laker go to green FIVE now		$n = 3$, No
8		Ready Tiger go to red EIGHT now		$n = 1$, No
9			Ready Tiger go to green SEVEN now	$n = 6$, Yes
10	Ready Arrow go to blue FOUR now			$n = 4$, No

This example shows the variable nontarget and color words used in the medium and high uncertainty conditions. In the medium-uncertainty condition, each of three talkers is associated with one of the three locations. In the high-uncertainty condition, the locations of the three talkers vary randomly across trials. A single talker is used with the same call sign and color on every trial in the minimal-uncertainty condition.

words, but had a consistent mapping of talker and location. This medium uncertainty condition creates the impression of a different specific person at each location, while maintaining the same amount of stimulus variability as in the high uncertainty condition. The factorial combination of these three conditions with the two Target conditions (call sign and number) resulted in six conditions. There were eight trial blocks in each condition, for a total of $30 \times 8 = 240$ observations per condition. Subjects were not told about the differences across conditions in the number of talkers, sentence variability, or the assignment of talkers to locations.

All subjects were presented with all six conditions, with a different counterbalanced order of conditions for each subject. Trial blocks were run in sets of four with no experimenter intervention between trial blocks within a set. All trial blocks within a set were in the same Target condition. The experimenter announced the identity of the target word at the beginning of each set and a reminder of the current target ("call sign" or "number") was displayed at the top of the screen throughout each trial block. The Target condition changed with each successive set, and the Uncertainty condition was held constant for two consecutive sets (one in each of the two Target conditions). Each counterbalanced order was created by using one of six possible orders of the three Uncertainty conditions and alternating Target conditions within each Uncertainty condition, starting with either call sign or number as target. One set of four trial blocks in each condition was run in the first test session, followed by a second set of four trial blocks in each condition in the second session, using the same order of conditions in each session. Testing was completed in two 90-min sessions on separate days.

Results

Response time (RT) and response correctness were scored on each trial. Response time was measured from the appearance of the "Yes" and "No" virtual buttons on the screen to the mouse click (or touch) on a button. Only RTs for correct responses were used in the analysis. Extreme fast and slow responses were omitted by excluding all RTs less than 150 ms and all RTs greater than three times the standard deviation above the mean for each condition. Using these exclusion criteria, the average number of excluded responses across conditions was less than three percent (almost entirely due to slow responses). For the purposes of statistical analysis, the percent-correct (PC) scores were converted to rationalized arcsine units (RAU; Studebaker, 1985).

Overall, performance was very good, with PC scores ranging from 80 to 96% (mean = 89%, $SD = 5.4\%$) for the YNH listeners and from 62 to 94% (mean = 78%, $SD = 9.3\%$) for the ONH listeners. Response times were similar to those found for other versions of the *n*-back task for the younger listeners (mean $RT =$ 780 ms, $SD = 295$ ms), but RTs for the older listeners (mean = 893 ms, $SD = 258$ ms) were more similar to the younger listeners than typically observed (see Verhaeghen and Basak, 2005; Basak and Verhaeghen, 2011).

The main results are summarized in **Figure 1**. Performance is shown as a function of *n* for both groups, with RT shown in **Figure 1A** and transformed percent correct (tPC) in RAU shown in **Figure 1B**. (Recall that *n* is the number back and the set size is constant at 3.) A 2 (Group) × 3 (Uncertainty) × 2 (Target) × 5 (*n* back) analysis of variance performed for both tPC and RT revealed that the group difference in RT was not significant ($F < 1.0$), while the difference in accuracy was significant [$F_{(1, 20)} = 9.84$; $p < 0.01$, $\eta_p^2 = 0.33$]. There were no interactions with group in either analysis ($p > 0.05$). Thus, both younger and older listeners with normal hearing were found to be affected by the experimental manipulations in the same way, with younger listeners significantly outperforming the older ones only in terms of accuracy. Because there were no interactions with the group variable, discussion of the effects of the within-group variables are presented below without a separate analysis for each group, although group-specific data will continue to be depicted descriptively in subsequent figures.

Performance as a Function of *n*

It can be seen in **Figure 1** that the effect of *n* was quite similar for the two groups for both RT and accuracy: RT tends to

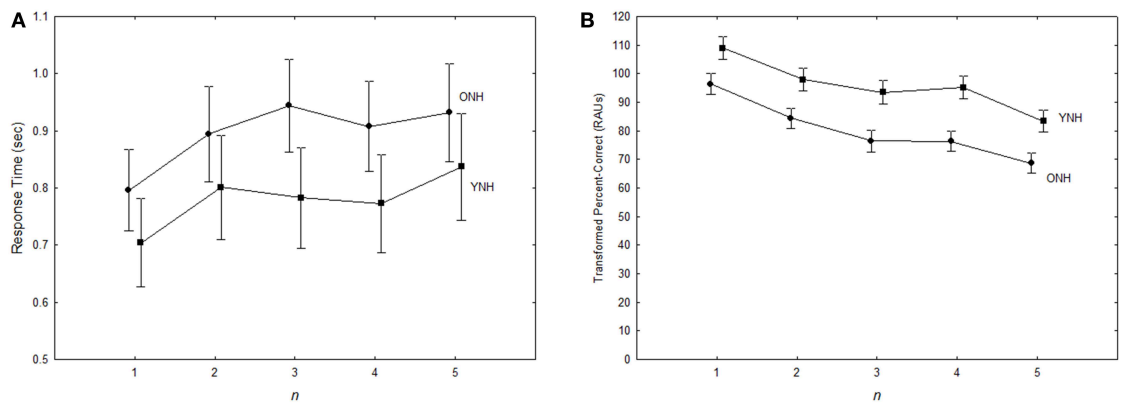


FIGURE 1 | Response times for correct responses (A) and accuracy (B) as a function of the number of presentations separating the target words to be compared, for both age groups in Experiment 1. Error bars indicate ± one standard error.

rise and accuracy tends to fall as n increases. Analysis of RTs revealed a significant main effect of n [$F_{(4, 80)} = 11.1$, $p < 0.001$,$\eta_p^2 = 0.36$]. However, follow-up analyses revealed that only the difference between $n = 1$ and $n = 2$ was significant (Tukey HSD; $p < 0.001$), with no significant differences for any further increases ($p > 0.25$). This indicates that both groups have the same cost (about 100 ms) for switching information in and out of working memory, and the same efficiency of memory search for items in the outer store. A significant n x Uncertainty interaction [$F_{(8, 160)} = 2.9$, $p < 0.01$, $\eta_p^2 = 0.13$] reflected a slight flattening of the RT function with an increase in uncertainty, with a significant difference between $n = 1$ and $n = 2$ only in the low-uncertainty condition (Tukey HSD, $p < 0.001$). This suggests that increased complexity of full sentences and irrelevant stimulus variability make it more difficult to access the most recent target word in memory. A significant three-way interaction [$F_{(8, 160)} = 2.3$, $p < 0.05$, $\eta_p^2 = 0.10$] reflected a larger performance decrement in the high-uncertainty condition with the call-sign target, especially for the lower values of n.

There was also a significant main effect of n for tPC scores [$F_{(4, 80)} = 61.3$, $p < 0.001$, $\eta_p^2 = 0.75$], with a negative accuracy slope of approximately 4 RAU. Each increase in n resulted in a significant decrease in tPC (Tukey HSD, $p < 0.05$), except for the difference between $n = 3$ and $n = 4$ ($p > 0.9$). The fairly constant difference between the two groups at all values of n shows that an increase in the time (and number of intervening items) between items to be compared resulted in similar decreases in the availability of items for young and older listeners.

Significant two-way interactions reflected slight differences in the rate of decrease in accuracy with increases in n in the different conditions. A significant Target x n interaction [$F_{(4, 80)} = 4.1$, $p < 0.005$, $\eta_p^2 = 0.17$] was associated with a substantially greater difference between performance for $n = 4$ and $n = 5$ for the call sign target than for the number target, and a significant Uncertainty x n interaction [$F_{(8, 160)} = 3.7$, $p < 0.001$, $\eta_p^2 = 0.16$] was due to a considerably smaller difference between $n = 1$ and $n = 2$ in the high uncertainty condition than in the other uncertainty conditions. Finally, a significant three-way interaction was primarily due to the latter two-way interaction being greater for the call sign target than for the number target.

Performance Under Different Levels of Uncertainty

Figure 2 shows the effect of uncertainty for YNH and ONH subjects for both RT and accuracy. Although performance was worst in the high-uncertainty condition for both measures, the pattern was slightly different for RT and tPC. Both main effects of uncertainty were significant [RT: $F_{(2, 40)} = 12.3$, $p < 0.001$, $\eta_p^2 = 0.38$; tPC: $F_{(2, 40)} = 4.5$, $p < 0.05$, $\eta_p^2 = 0.18$] and follow-up tests (Tukey HSD) indicated a similar pattern of significance for both RT and tPC. For RT, the high-uncertainty condition was significantly slower than the other uncertainty conditions ($p < 0.001$), which were not different from each other ($p > 0.9$). For tPC, the low- and medium-uncertainty conditions were not significantly different ($p > 0.6$), and the high-uncertainty condition was significantly more difficult than the low-uncertainty condition ($p < 0.05$). However, the difference between the high- and medium-uncertainty conditions was only marginally significant ($p < 0.1$). Thus, the advantage of the constant mapping of voice and location in the medium-uncertainty condition was more robust in terms of RT than accuracy. For both measures, the ecological validity of the constant mapping in the medium-uncertainty condition led to better performance, equal to that in the low-uncertainty condition, despite having the same degree of stimulus variability across trials as in the high-uncertainty condition.

Performance with Early and Late Target Words

Figure 3 shows performance as a function of the target word for both groups. It can be seen that both YNH and ONH subjects were consistently slower [$F_{(1, 20)} = 38.1$, $p < 0.001$, $\eta_p^2 = 0.66$],

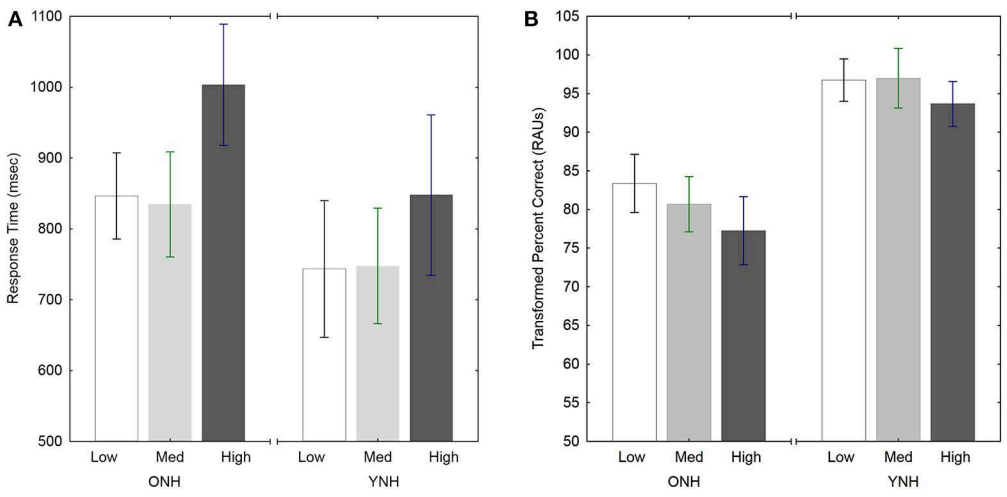


FIGURE 2 | Response times (A) and accuracy (B) for both groups of subjects in the three Uncertainty conditions of Experiment 1. Error bars indicate ± one standard error.

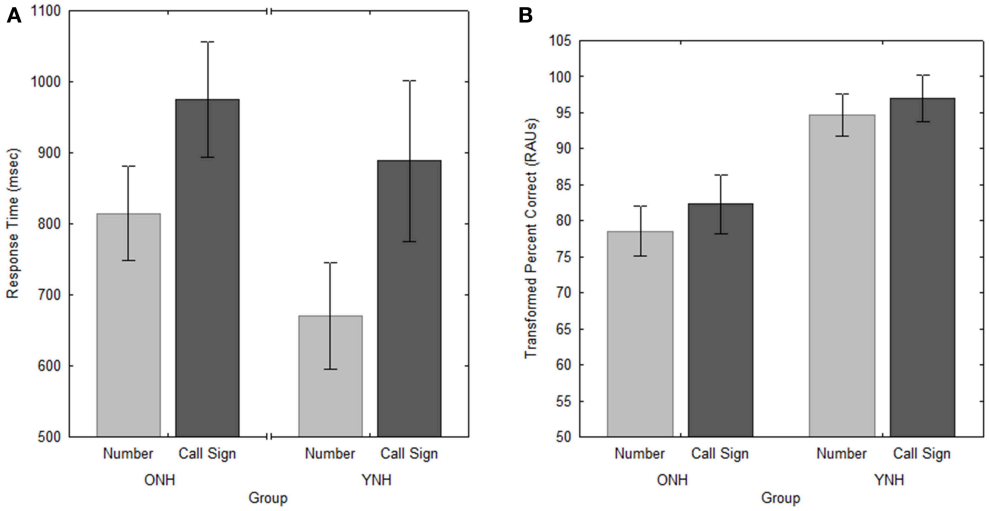


FIGURE 3 | Response times (A) and accuracy (B) for both groups of subjects in the two Target conditions in Experiment 1. Error bars indicate ± one standard error.

but slightly more accurate [$F_{(1,\ 20)} = 7.5, p < 0.05, \eta_p^2 = 0.27$], when responding to the call sign than to the number target. Because the call sign occurred early in each sentence, subjects had more time to prepare their response before the "yes" and "no" response buttons appeared (and the RT timer started) at the end of the sentence presentation. That subjects were unable to use this time to decrease RT suggests that the irrelevant words following the call sign may have interfered with memory or decision processes. The slightly more accurate responding may be due to the greater distinctiveness for call signs (highly distinguishable two-syllable names), compared to numbers, which were more similar single-syllable (with the exception of "seven") numerals.

Discussion

In addition to providing measures of working-memory abilities, this modified version of the *n*-back task, using full sentences from the CRM corpus, was designed to assess recall abilities using a listening situation that had some features in common with a natural sequential multitalker conversation. In many ways, performance on this task was similar to that obtained with versions of the *n*-back task that used much simpler visual stimuli and similar strategies for varying *n* (e.g., Verhaeghen and Basak, 2005; Oberauer, 2006; Basak and Verhaeghen, 2011). Both younger and older subjects showed a significant switching cost as evidenced by an increase in RT as *n* (the number back) increased from 1 to 2, and neither group showed any further increases in RT as *n* increased from 2 to 5. It is important to remember that *n* in the present study is not equal to the set size, as is common in *n*-back studies. Because set size is held constant here at 3 (the number of locations), any increase in RT with an increase in *n* would be attributed to an increase in the time between comparison items rather than to an increase in the number of items in a search set. The results also agreed with the earlier visual *n*-back studies in showing no age differences in the switching cost. However, in contrast to the earlier studies, no correction for general slowing was required, because RTs were very similar for younger and older subjects. Thus, not only were there no age differences in accessibility of items in the focus of attention, there was little or no evidence of slowing of memory retrieval or decision making with age in this task.

On the other hand, age differences *were* observed with accuracy in the present task. Older subjects were consistently less accurate, by about 10 percentage points, than younger subjects for all values of *n*. In the related visual *n*-back studies, age differences were not found for $n = 1$ when set size was confounded with the number back, but, when they were not confounded, as in the present experiment, age differences were also found for all values of *n* (Verhaeghen and Basak, 2005; Basak and Verhaeghen, 2011). Thus, older subjects appear to have more trouble maintaining an item in memory, whether it is in the focus of attention or in the outer store, at least under some task conditions. Despite this, when older subjects correctly recalled the repetition of the current item (or lack of it) in a given location, they were not significantly slower than younger subjects in recall and decision making. Thus, aging appears to affect the ability to hold information in memory in this task, but not the ability to access and make judgments on that information when the information is available.

Variability in talkers and nontarget words across trials in this task had a detrimental effect on performance for both younger and older subjects. Subjects were fastest and most accurate when the talker and nontarget words were held constant across trials (low uncertainty) and slowest and least accurate when those words varied randomly (high uncertainty). However, when talkers were assigned to unique locations (as they typically are in most conversational settings), performance was just as good as in the low-uncertainty condition, despite the same amount of talker and semantic variability as in the high-uncertainty condition. This shows that both older and younger listeners are sensitive to location information and voice information, and that

a consistent mapping of these two types of information is helpful when trying to keep track of what was said in a sequence of spoken sentences. In a sense, this mapping can be thought of as a reduction of uncertainty, in that subjects know what voice to expect from each location. But because subjects cannot predict which location will be used on the next trial and cannot identify the location or the talker until after the sentence begins, it seems unlikely that the consistent mapping advantage is simply due to reduced variability in the mapping of stimulus properties that are irrelevant to the task. It seems more likely that the variable talker-location mapping adversely affects performance because it is a violation of an expectation based on everyday experience.

The findings with regard to selection of the early-occurring target (call sign) or late-occurring target (number) are difficult to interpret. Both groups were substantially slower in making judgments about the repetition of the call sign, but they were slightly more accurate than with the number target. While this is consistent with a speed-accuracy tradeoff, the significant, but rather small, increase in transformed percent-correct performance (less than 2 RAU) may not entirely account for the relatively large increase in RT (nearly 200 ms, or about 26%) and it is not clear why subjects would use a different speed-accuracy tradeoff based on the target word. Despite giving subjects more time to prepare their response before the sentence ended (and the RT clock started), the greater time and number of intervening words between the early target word and the presentation of the response options appears to have made it more difficult for subjects to access the item in memory. It thus appears that this retroactive interference slowed recall and/or decision processes without affecting the availability of the target word in memory.

Experiment 2: Older Hearing-impaired Adults with and without Amplification

The older subjects in Experiment 1 generally performed well on the auditory *n*-back task, but although they were about as fast as the younger subjects in all conditions, they were consistently less accurate. This pattern of results suggests that, at least with the present task, aging affects the ability to hold information in working memory while processing new information, but not the ability to access information in working memory and make rapid judgments based on it. However, the older subjects in Experiment 1 had relatively good hearing and showed no evidence of having any difficulty understanding the talkers. Because hearing loss is common in the older population, it is important to determine whether older listeners with poorer hearing perform differently on this type of auditory memory task. If listeners have to expend more effort trying to understand what is being said, they may be more susceptible to memory interference and uncertainty in ways that lead to a different pattern of results from that observed in Experiment 1 (see McCoy et al., 2005; Pichora-Fuller and Singh, 2006; Gosselin and Gagné, 2011; Rudner et al., 2012; Yusuf et al., 2012).

To examine the effect of hearing loss on performance with this task, Experiment 2 employed a group of hearing-impaired subjects who performed the auditory *n*-back task with and without custom spectral shaping (amplification) to ensure audibility of the speech materials. It was expected that without spectral shaping, the added difficulty would cause these listeners to be: (1) slower than their normal-hearing age peers; (2) more affected by stimulus uncertainty; (3) less able to take advantage of location and voice cues, and thus less able to take advantage of a constant talker-location mapping; and (4) more affected by target position because of a greater susceptibility to interference from irrelevant words following the early target. With shaping, these listeners were expected to be more like the older normal-hearing listeners. However, because this group may suffer from cochlear pathology and may have undergone changes in higher-level processing, either central auditory or cognitive processing (Humes et al., 2012), they were not expected to perform the same as the ONH listeners in Experiment 1.

Methods

Subjects

The subjects in this experiment were 11 older hearing-impaired listeners whose ages ranged from 64 to 85 years (mean = 70.1 years; $SD = 5.7$). There were five females and six males; two were current hearing aid users, and the others had never worn hearing aids. The highest level of education completed ranged from high school (one subject) to vocational school (two subjects), college (four subjects), and graduate school (four subjects). All subjects had symmetrical high-frequency sensorineural hearing loss and failed to meet the definition of normal hearing used in Experiment 1 (as described above). Thresholds for all subjects are shown in **Figure 4**. Except for hearing thresholds, the inclusion criteria were the same as for the older subjects in Experiment 1, and all had previously participated in the same individual differences study by Humes et al. (2013) as had the ONH subjects in Experiment 1.

Stimuli

The stimuli were the same CRM sentences used in Experiment 1, presented with and without custom amplification to ensure audibility. In the unshaped condition, the same 85-dB SPL level used in Experiment 1 was used in this experiment. In the shaped condition, presentation levels were adjusted to ensure that speech information was audible and to provide comparable presentation levels for all listeners. The levels were adjusted by measuring the long-term spectrum of the full set of stimuli and filtering each stimulus to shape the spectrum according to each listener's audiogram. The shaping was applied with a 68 dB SPL overall unshaped speech level as the starting point, and gain was applied as necessary at each 1/3 octave band to produce speech presentation levels at least 13 dB above threshold from 125 Hz to 4000 Hz.

Procedures

Testing procedures were the same as in Experiment 1, using the same equipment. All subjects were tested twice: once with shaping and once without shaping, each time following the same procedures and including all the conditions described for Experiment 1. Based on a random assignment, five subjects were

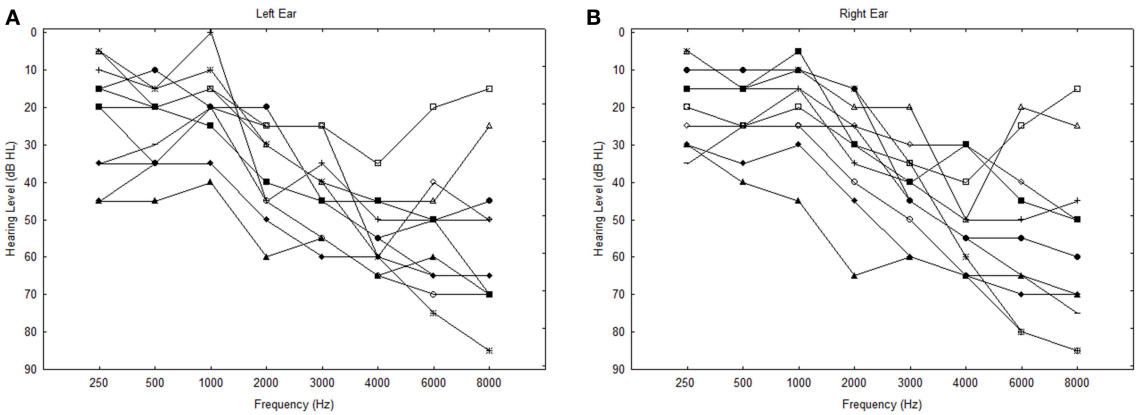


FIGURE 4 | Hearing thresholds for the left (A) and right (B) ears of the OHI subjects in Experiment 2.

tested with unshaped stimuli first, and six were tested with shaped stimuli first. Testing was completed in four 90-min sessions on four separate days.

At the end of the experiment, a short recognition test was conducted to determine whether subjects were able to understand the words in the CRM sentences at the levels used in the experiment. The sentences were presented both with and without shaping, using the right ear only. Subjects listened to the same CRM sentences used in the main experiment (using the same talkers) and indicated the call sign, color, and number in each sentence by touching (or clicking with a mouse) virtual buttons on the monitor labeled with all of the possible options for each of the three target words. There were 16 blocks of 32 trials: 8 blocks with shaping and 8 blocks without shaping, using the same counterbalanced order of shaping conditions used in the main experiment.

Results

On the post-experiment recognition test, all subjects correctly identified all target words on every trial, clearly demonstrating that the stimuli were audible under the presentation conditions used in this experiment. Thus, the deviations from perfect performance described below must be attributed to the memory and processing requirements of the task.

Response times and accuracy were analyzed as in Experiment 1, using the same exclusion criteria for outliers in the RT data and resulting in similar rejection rates. Overall, subjects' accuracy was very close to the 78% correct obtained with the ONH subjects in Experiment 1, with 80% correct overall for both shaped and unshaped testing. However, RTs were considerably slower. Average RTs across all conditions were 1561 ms ($SD = 544$ ms) without shaping and 1475 ms ($SD = 618$ ms) with shaping, a nearly 70% increase relative to the ONH subjects in Experiment 1. The slow mean response time for this older group was partly due to one listener (the oldest, at 85 years) whose average RT was about 2.6 standard deviations above the group mean. (This subject was retained because performance was above chance and response times showed systematic variation with conditions.) However, even without this subject, mean performance was still 470 ms slower than for the ONH subjects in Experiment 1. This difference was statistically significant whether evaluated with or without the slowest subject [$t_{(20)} = 3.26$, $p < 0.005$ and $t_{(19)} = 3.79$, $p < 0.005$, respectively].

Analysis of variance was performed, using a 2 (shaping/no shaping) by 3 (Uncertainty) × 2 (Target) × 5 (*n*-back) design for both RT and percent-correct performance (RAU transformed). No effect of shaping was observed for either RT or tPC (Fs < 1.0), and there were there no interactions with shaping for either measure ($p > 0.05$). As in Experiment 1, there was a significant effect of *n* for both RT [$F_{(4, 40)} = 7.74$, $p < 0.001$, $\eta_p^2 = 0.44$] and tPC [$F_{(4, 40)} = 53.60$, $p < 0.001$, $\eta_p^2 = 0.84$], as well as significant effects of Target [for RT, $F_{(1,10)} = 7.60$, $p < 0.05$, $\eta_p^2 = 0.43$; for tPC, $F_{(1,10)} = 16.63$, $p < 0.005$, $\eta_p^2 = 0.62$], but the Uncertainty manipulation did not have a significant effect in this Experiment ($p > 0.05$ for both RT and PC).

The main results are summarized in **Figure 5**, which shows RT as a function of the number back (*n*) in **Figure 5A**, and tPC vs. *n* in **Figure 5B**. The pattern of performance for both RT and tPC was essentially the same as in Experiment 1. There was a clear cost of switching information in and out of the focus of attention, as seen by the increase in RT between $n = 1$ and $n = 2$ (Tukey HSD, $p < 0.01$), with no significant changes in RT with further increases in *n* ($p > 0.05$). Also as in Experiment 1, the decrease in tPC with *n* was significant for successive increases in *n* (Tukey HSD, $p < 0.01$), except for that between $n = 3$ and $n = 4$ ($p > 0.05$).

A significant Uncertainty by *n*-back interaction [$F_{(8, 80)} = 2.8$, $p < 0.01$, $\eta_p^2 = 0.22$] in the RT data was primarily due to a reduced switching cost for the high-uncertainty condition. This was the only Uncertainty condition in which performance was not consistently better for $n = 1$ than for $n > 1$, with RT for $n = 1$ not significantly better than for $n = 3$ or $n = 5$ (Tukey HSD, $p > 0.05$). A significant 3-way interaction between Target, Uncertainty, and *n*-back [$F_{(8, 80)} = 3.3$, $p < 0.01$, $\eta_p^2 = 0.25$] reflected the fact that this reduced switching cost was greater for the number than for the call-sign target.

There were two significant interactions in the tPC data. A Target by *n*-back interaction [$F_{(4, 40)} = 6.7$, $p < 0.001$, $\eta_p^2 =$

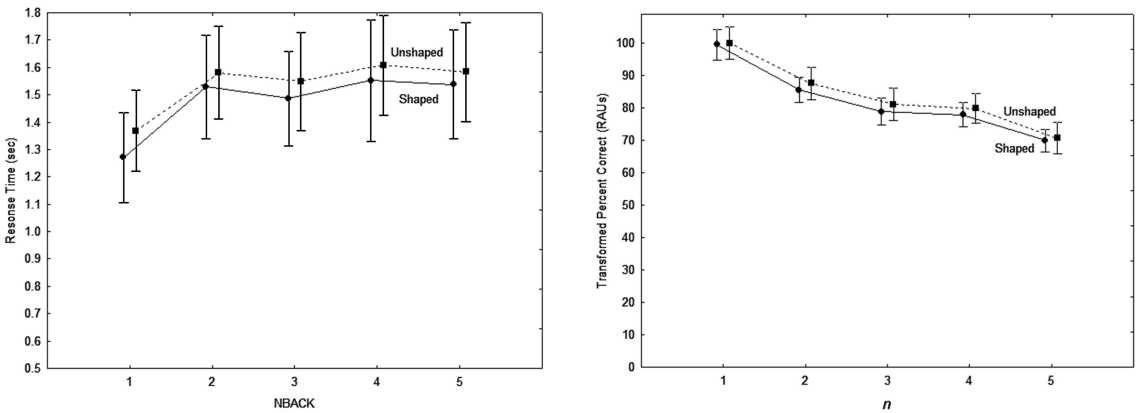


FIGURE 5 | Response times for correct responses (A) and accuracy (B) as a function of the number of presentations separating the target words to be compared, for OHI subjects with and without shaping in Experiment 2. Error bars indicate ± one standard error.

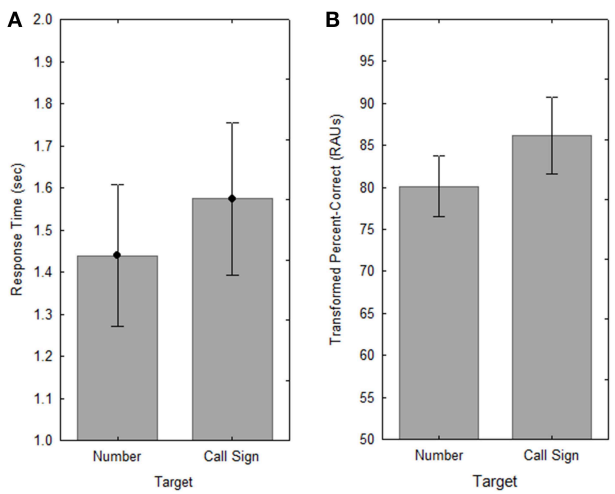


FIGURE 6 | Response times (A) and accuracy (B) for the two Target conditions in Experiment 2. Error bars indicate ± one standard error.

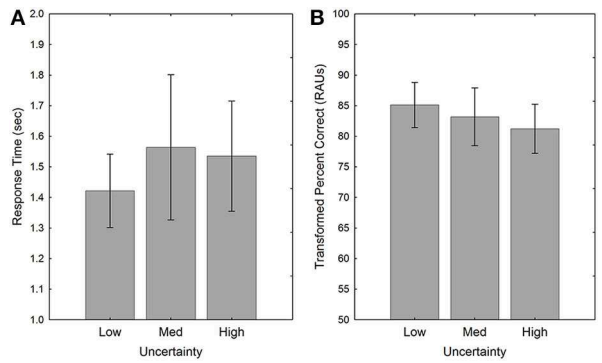


FIGURE 7 | Response times (A) and accuracy (B) for the three Uncertainty conditions in Experiment 2. Error bars indicate ± one standard error.

0.40] was due to the lack of a significant effect of Target for $n = 1$ or for $n = 5$ (Tukey HSD, $p > 0.5$). An Uncertainty by n-back interaction [$F_{(8, 80)} = 2.3$, $p < 0.05$, $\eta_p^2 = 0.19$] reflected a tendency for greater differences between the three uncertainty conditions for $n = 1$ and $n = 5$ than for other values of n.

The effect of Target was essentially the same as in Experiment 1, with significantly slower RT [$F_{(1, 10)} = 7.6$, $p < 0.05$, $\eta_p^2 = 0.43$] and greater accuracy [$F_{(1, 10)} = 16.6$, $p < 0.005$, $\eta_p^2 = 0.62$] for judgments of repetition of the call sign in a given location than for repetitions of the number (see **Figure 6**). This is suggestive of the same speed-accuracy tradeoff seen in Experiment 1, although the increase in RT for the call sign was slightly smaller than that seen with the older listeners in Experiment 1 (approximately 135 ms; a 9% increase) and the corresponding change in PC of roughly 4 percentage points was slightly higher.

Although the effects of target identity (or sentence position) and the number of intervening sentences between to-be-compared items were quite similar to those observed in Experiment 1, this was not the case with the uncertainty manipulation. Although there was a slight tendency for RT to increase and for accuracy to decrease as the level of uncertainty increased (see **Figure 7**), these differences were not significant, and there was no evidence of an advantage for the consistent mapping of talker and location, as observed in Experiment 1.

General Discussion

These experiments used a modified n-back task with auditory presentation of sentences to examine the effects of aging and hearing loss on the ability to understand and remember spoken material and to keep track of source locations. By asking subjects to compare a target word in a sentence just heard to the corresponding target word in the last sentence presented from the same location (left, center, or right), the task eliminates the need

to keep track of the number of trials between comparison items, as is commonly required with the *n*-back task. This makes the task more natural, and when a specific talker is associated with a specific location, the task becomes similar to keeping track of who said what in a typical conversational setting.

In many respects, the pattern of results was similar to that from earlier studies using visual presentation of digits in which the to-be-compared items were also identified by location (a column in a visual display), rather than by a fixed number back in a series of presentations (e.g., Verhaeghen and Basak, 2005; Basak and Verhaeghen, 2011). The added complexity of full sentences, rather than single digits, did not change the basic pattern of response times and accuracy as a function of the number back (*n*). There was a similar cost of switching items in and out of the focus of attention, as evidenced by the increase in RT between $n = 1$ and $n = 2$, and no further increases in RT as *n* increased beyond 2. However, unlike the earlier visual studies, the set size was held constant, and changes in *n* were associated only with longer delays between items to be compared. In the earlier studies, set size was varied and results were plotted as a function of set size whether it was confounded with the number back (as in Verhaeghen and Basak, 2005) or varied independently from the number back (as in Oberauer, 2006; Basak and Verhaeghen, 2011). Basak and Verhaeghen did not examine the effect of the number back, partly because the focus was on set size (which was equal to the number of positions), but also because of the constraint that all positions must be tested before any position is repeated. With this constraint, the position to be tested on a given trial becomes more predictable as the number of untested positions in the set decreases. However, Oberauer (2006) did not use the same position-sampling constraint and could thus examine the effect of the number back (or *lag*) independently of set size. He found that both RT and PC were affected by lag as well as by set size, with linear increases in RT and linear decreases in accuracy as the lag increased. There was no evidence of the flattening of the RT function after $n = 2$, as observed in the present study.

Given the use of sentences in the present study, which introduce longer presentation times (and intertrial intervals) and a greater potential for interference than single digits, it is perhaps surprising that response times did not increase when the number of intervening sentences between compared target words increased. However, accuracy did decrease linearly with *n*, suggesting that memory interference and decay were occurring with time and number of stimuli presented. The lack of an effect of the number of intervening sentences on response times for correct responses after $n = 2$ (when a focus switch was required) indicates that when the information is available in memory, access time and decision time are not slowed. Thus, it is primarily the likelihood of a correct response (or the availability of information) that decreases with *n*, not the accessibility of the information stored in memory.

The present findings provide no evidence to suggest that the effect of the number back on response times or accuracy changes with age or hearing loss. Although the older hearing-impaired listeners in Experiment 2 were substantially slower than those in Experiment 1, they showed roughly the same switching cost and no further increases in response times with increasing values of *n*. Moreover, there was no main effect of age on response times in Experiment 1 and there were no interactions involving age. The only effect of age was on accuracy, but there were no interactions involving age for the accuracy measure either. Older subjects were less accurate than younger ones, but this did not vary with the number back or any other experimental manipulation in Experiment 1. Thus, it appears that aging primarily affects the susceptibility to decay and interference of information inside and outside of the focus of attention, while having little or no effect on the accessibility of information that is retained.

Age differences were also absent in the effect of target word location. It was expected that older subjects might have more trouble with the early (call sign) targets because of a greater susceptibility to interference from the following words in the sentence. In this task, a judgment about the repetition of a target word can be made as soon as the word is recognized, but the response cannot be made until the end of the sentence, when the response options are presented (and the response timer starts). Thus, faster responding would be expected for early targets if subjects could make their judgments early and prepare their response while ignoring the rest of the sentence. However, both younger and older subjects were unable to take advantage of this, responding more slowly to early targets than to later (number) targets, despite being slightly more accurate with the early targets. Although it is not clear why responses were slower for the early targets, both groups appear to have required more time for recall and decisions regarding the early targets, even though the information was at least as available (as indicated by accuracy scores) as it was for the later targets. It may be that interference or distraction from words following the target word make younger and older listeners less confident in their responses, thus slowing response times without affecting accuracy.

Although the OHI subjects in Experiment 2 showed the same effect of *n* and target word on response times and accuracy, they had much slower response times than the ONH subjects in Experiment 1, with a mean difference of more than 600 ms (roughly 1.7 times greater). The difference was fairly consistent across subjects; only three subjects in Experiment 2 had mean RTs below 1 sec, while all but 3 of the 11 ONH subjects in Experiment 1 had RTs below 1 sec. Shaping, to ensure audibility, had no effect on response times or accuracy, and none of the effects in Experiment 2 were impacted by the shaping manipulation. The slower response times do not appear to be due to an inability to reliably understand the target words, because subjects were as accurate as the ONH subjects in Experiment 1 even without shaping, and they performed perfectly on a target-word recognition test using the same stimulus materials presented at the same levels used in Experiment 2. Although the average age for the OHI group was about 5 years greater than that for the ONH group, age was not significantly correlated with RT. The oldest subject (85 years) was the slowest by a large margin, but, with this extreme subject excluded, the correlation between age and RT for all HI subjects was 0.05. Finally, even cognitive abilities, as measured by a global cognitive ability factor

obtained in an earlier study (Humes et al., 2013), do not account for the slower response times. The extremely slow subject did score quite poorly on the cognitive measure (based on three working memory measures and a processing speed test), but the correlation between that measure and RT was small and non-significant, with ($r = -0.26$) or without ($r = 0.14$) the extreme subject included.

Another potential explanation for the substantially slower RTs in Experiment 2 is that these hearing-impaired subjects had to expend more effort to understand the spoken sentences than did the normal hearing listeners in Experiment 1. Subjects with mild to moderate hearing loss often must expend more effort than their normal-hearing peers to achieve comparable levels of speech understanding, and this is not always evident in speech-recognition performance (see Rabbitt, 1991; Pichora-Fuller et al., 1995; Tun and Wingfield, 1999; McCoy et al., 2005; Gosselin and Gagné, 2011). This research suggests that the emphasis on memory over immediate recognition in the current task would be expected to make it sensitive to differences in effort, especially in an older population, which is likely to have declining cognitive abilities. The shaping used in Experiment 2 should have reduced the amount of effort required by reducing reliance on partial information to compensate for inaudible portions of speech. However, the lack of an effect of shaping does not rule out an effort-based explanation for the slower response times in Experiment 2. Although the provided shaping ensures audibility from 125 to 4000 Hz, the listening experience is not equivalent to that for normal-hearing listeners. The listeners in Experiment 2 were not experienced hearing aid users (only two had ever worn hearing aids), and the amplified speech signal presented cannot be expected to provide a listening experience equivalent to normal hearing. Support for the effort explanation was also lacking in the correlations between hearing loss (PTA and HFPTA) and RT within this group of OHI listeners: the correlations were not significant and the tendency was in the wrong direction (with greater hearing loss associated with slightly faster response times). However, lack of an association between hearing loss and RT within a small group of hearing-impaired listeners is not strong evidence against the effort explanation.

It thus appears that the considerably slower response times of the OHI group may be due to an increase in the effort required to understand speech, which is commonly associated with hearing loss. That accuracy was essentially the same as for the ONH subjects in Experiment 1 indicates that the OHI subjects understood and retained the target words about as well as the ONH subjects. The longer response times thus indicate difficulty accessing the stored information, lower confidence in their judgments, or both. Although lower confidence is often associated with longer response times (e.g., Emmerich et al., 1972; Vickers and Packer, 1982), it is not possible to determine the relative contributions of access time and decision time to response latencies in the present study.

Uncertainty and the Use of Location and Voice Information

The use of different talkers and virtual spatial locations in this study allowed for an examination of the ability to use location and voice information in a speech-understanding task as a function of age and hearing loss. It also allowed for the introduction of greater stimulus variability across trials by varying location and talker as well as the words (target and nontarget) used in the CRM sentences. The Uncertainty variable in this study included three levels of stimulus variability, or uncertainty, that utilized two types of assignment of talkers to spatial locations: consistent and variable. The normal-hearing subjects in both age groups in Experiment 1 were affected the same way by the uncertainty manipulation. Responses were slower and less accurate with the highest level of uncertainty, when the voice, location, and nontarget words varied randomly over trials, than in the low-uncertainty condition, in which the same voice and nontarget words were used on every trial. However, in the medium-uncertainty condition, with consistent mapping of talkers to locations (but with the same amount of stimulus variability as the high-uncertainty condition), response times were roughly the same as in the minimal-uncertainty condition, and accuracy followed a similar pattern. Thus, the decline in performance across the three uncertainty conditions was almost entirely due to the difference between consistent and inconsistent mapping of voice and location. Although the use of a single talker in three locations in the minimal uncertainty condition is not a natural situation, this was offset by the lack of variability in voice and nontarget words. When there was variation in talkers (voices), the ecological validity of a consistent location for each talker eliminated the effect of the increased stimulus variability on response times and nearly so for accuracy. This suggests that it was the unpredictable change in talkers (not simply variation in the talker and the nontarget words) that was primarily responsible for the increased difficulty in the high-uncertainty condition.

In contrast to the normal-hearing listeners, the older hearing-impaired listeners in Experiment 2 were unable to take advantage of the consistent voice/location mapping in the medium-uncertainty condition. Although there were small differences between uncertainty conditions favoring minimal uncertainty, the effect of uncertainty was not significant in this group, and response times with consistent mapping were nearly identical to those with the inconsistent mapping of the high-uncertainty condition. Thus, despite being just as good in recognition and recall accuracy as the ONH subjects in Experiment 1, the OHI subjects did not find the predictability of a consistent mapping of voice and location information helpful. Given that the ability to discriminate the three virtual locations is required to perform this task, it is unlikely that localization problems were a significant factor. However, difficulty in reliable discrimination of the three voices may have been responsible for the failure to benefit from consistent mapping. Although the three talkers used in this study are highly discriminable for young normal-hearing listeners, the older hearing-impaired listeners may not have been as sensitive to the voice differences. However, given that older hearing-impaired listeners have been shown to be adversely affected by talker uncertainty in recognition tasks using these CRM stimuli (Humes et al., 2006; Humes and Coughlin, 2009), it is unlikely that poor talker discrimination abilities fully account for the lack of benefit in the consistent mapping condition. It

seems more likely that the same factors that cause OHI listeners to require more time to make memory-based judgments also reduce their sensitivity to more subtle stimulus characteristics that are not a necessary component of the task. That is, the reduction in available resources, due to the increased effort expended by OHI listeners when listening to the sentences, may also reduce their sensitivity to talker differences in the context of a multitalker listening task that emphasizes memory over recognition abilities.

Summary and Conclusions

This study used a modified auditory *n*-back task with multiple talkers and locations to approximate the demands of a sequential multitalker conversation. Young and older adult listeners, with and without hearing loss, were asked to judge whether a target word in a sentence just heard was the same as in the last sentence heard from a given location. Performance on this task was similar to that obtained in a comparable version of the *n*-back task in the visual modality. Younger and older subjects with normal hearing showed similar costs when switching information in and out of the focus of attention and had similar response times overall. Neither group showed any increase in response times with greater numbers of trials between comparison words when comparing target words outside of the focus of attention (i.e., for comparisons with words presented more than one trial back in the sequence). Age did have an effect on the accuracy of the judgments; both groups were less accurate as the interval between comparison words increased, but older subjects performed consistently worse for all intervals between comparison words, whether or not focus switching was required. Older subjects with hearing loss showed a similar pattern of results, but had considerably longer response times, despite responding as accurately as the older normal-hearing listeners. All subjects responded more slowly and slightly more accurately to early target words than to later target words, showing no evidence of differential interference with age or hearing loss from the greater number of irrelevant words following the early target word.

Normal-hearing listeners in both age groups showed essentially the same adverse effect of stimulus uncertainty, but performed better under high uncertainty when talkers were consistently assigned to specific locations, rather than varying randomly across trials. However, older hearing-impaired listeners, in addition to responding more slowly than older normal-hearing listeners, showed no effect of stimulus uncertainty and were not helped by the ecological validity of a consistent mapping of voice and location. The slower response times and insensitivity to consistent talker/location mapping for the older hearing-impaired listeners, despite accuracy equal to that for the normal-hearing older adults, suggest that the older hearing-impaired listeners may have exerted more effort to perform at the same level of accuracy. This may have led to slower response times (perhaps related to reduced confidence) and reduced sensitivity to voice characteristics that can be helpful in reducing talker uncertainty (when talker identity is predicted by the location) and facilitating target word comparisons when the words are spoken by the same talker. It should be noted that these effort-based effects were observed using presentation levels well above typical conversational levels, and with customized spectral shaping. It is likely these effects would have been greater if the stimuli were presented at normal conversational levels.

These findings show that when simple speech-recognition tasks are complicated by memory requirements that begin to resemble the demands of a typical sequential multitalker conversation, hearing impairment, especially when combined with aging, can make it more difficult to keep track of what has been said and by whom. Although hearing loss primarily affected response times rather than accuracy in the present study, slower response times may result from greater effort, which can cause fatigue and reduce accuracy after prolonged periods of listening, especially under more difficult listening conditions. Moreover, a reduction in attentional resources that results in reduced sensitivity to voice characteristics may also diminish a listener's ability to notice other indexical properties or prosodic information that can be critical for effective communication.

Acknowledgments

The authors thank Kristen Baisley, Tera Quigley, Hannah Fehlberg, and Megan Chaney for assistance with various aspects of data collection for this study, and Patricia Knapp for editorial assistance. This work was supported, in part, by a research grant from the National Institute on Aging (R01 AG008293).

References

ANSI. (1999). *Maximum Permissible Ambient Levels for Audiometric Test Rooms (S3.1-1999).* New York, NY: American National Standards Institute.

ANSI. (2004). *S3.6-2004, Specification for Audiometers*. New York, NY: American National Standards Institute.

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Basak, C., and Verhaeghen, P. (2011). Aging and switching the focus of attention in working memory: age differences in item availability but not in item accessibility. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 66B, 519–526. doi: 10.1093/geronb/gbr028

Bolia, R. S., Nelson, W. T., Ericson, M. A., and Simpson, B. D. (2000). A speech corpus for multitalker communications research. *J. Acoust. Soc. Am.* 107, 1065–1066. doi: 10.1121/1.428288

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychon. Bull. Rev.* 12, 769–786. doi: 10.3758/BF03196772

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185. doi: 10.1017/S0140525X01003922

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433.

Dobreva, M. S., O'Neill, W. E., and Paige, G. D. (2011). Influence of aging on human sound localization. *J. Neurophysiol.* 105, 2471–2486. doi: 10.1152/jn.00951.2010

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2005a). Word recognition in noise at higher-than-normal levels: decreases in scores and increases in masking. *J. Acoust. Soc. Am.* 118, 914–922. doi: 10.1121/1.1953107

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2005b). Recognition of filtered words in noise at higher-than-normal levels: decreases in scores with and without increases in masking. *J. Acoust. Soc. Am.* 118, 923–933. doi: 10.1121/1.1953127

Emmerich, D. S., Gray, J. L., Watson, C. S., and Tanis, D. C. (1972). Response latency, confidence, and ROCs in auditory signal detection. *Percept. Psychophys.* 11, 65–72.

Gosselin, P. A., and Gagné, J.-P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *J. Speech Lang. Hear. Res.* 54, 944–958. doi: 10.1044/1092-4388(2010/10-0069)

Houtgast, T., and Festen, J. M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *Int. J. Audiol.* 47, 287–295. doi: 10.1080/14992020802127109

Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *J. Acoust. Soc. Am.* 112, 1112–1132. doi: 10.1121/1.1499132

Humes, L. E., and Coughlin, M. P. (2009). Aided speech-identification performance in single-talker competition by elderly hearing-impaired listeners. *Scand. J. Psychol.* 50, 485–494. doi: 10.1111/j.1467-9450.2009.00740.x

Humes, L. E., and Dubno, J. R. (2010). "Factors affecting speech understanding in older adults," in *The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis, Springer Handbook of Auditory Research (SHAR)*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R.Fay (New York, NY: Springer-Verlag), 211–258.

Humes, L. E., Dubno, J. R., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J., et al. (2012). Central presbycusis: a review and evaluation of the evidence. *J. Am. Acad. Audiol.* 23, 635–666. doi: 10.3766/jaaa.23.8.5

Humes, L. E., Kidd, G. R., and Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front. Syst. Neurosci.* 7:55. doi: 10.3389/fnsys.2013.00055

Humes, L. E., Lee, J. H., and Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *J. Acoust. Soc. Am.* 120, 2926–2937. doi: 10.1121/1.2354070

Kane, M. J., Conway, A. R. A., Miura, T. K., and Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: a question of construct validity. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 615–622. doi: 10.1037/0278-7393.33.3.615

Kidd, G. R., and Humes, L. E. (2012). Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences. *J. Acoust. Soc. Am.* 131, 1434–1448. doi: 10.1121/1.3675975

Martin, R. L., Flanagan, P., McAnally, K. I., and Eberle, G. (2011). Memory for the locations of environmental sounds. *J. Acoust. Soc. Am.* 129, 3873–3883. doi: 10.1121/1.3589254

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* 58, 22–33. doi: 10.1080/02724980443000151

McElree, B. (2001). Working memory and focal attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 817–835. doi: 10.1037/0278-7393.27.3.817

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *J. Exp, Psychol. Learn. Mem. Cogn.* 28, 411–421. doi: 10.1037/0278-7393.28.3.411

Oberauer, K. (2006). Is the focus of attention in working memory expanded through practice? *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 197–214. doi: 10.1037/0278-7393.32.2.197

Oberauer, K., and Bialkova, S. (2009). Accessing information in working memory: can the focus of attention grasp two elements at the same time? *J. Exp. Psychol. Gen.* 138, 64–87. doi: 10.1037/a0014738

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm.20131

Parmentier, F. B. R., and Jones, D. M. (2000). Functional characteristics of auditory temporal-spatial short-term memory: evidence from serial order effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 222–238. doi: 10.1037/0278-7393.26.1.222

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Rabbitt, P. M. A. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol. Suppl.* 476, 167–176. doi: 10.3109/00016489109127274

Redick, T. S., and Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: a meta-analysis. *Psychon. Bull. Rev.* 20, 1102–1113. doi: 10.3758/s13423-013-0453-9

Rudner, M., Lunner, T., Behrens, T., Thoren, E. S., and Ronnberg, J. (2012). Working memory capacity may influence perceived effort during aided speech recognition in noise. *J. Am. Acad. Audiol.* 23, 577–589. doi: 10.3766/jaaa.23.7.7

Salthouse, T. A. (2010). *Major Issues in Cognitive Aging*. New York, NY: Oxford University Press.

Salthouse, T. A. (2012). Consequences of age-related cognitive declines. *Annu. Rev. Psychol.* 63, 201–226. doi: 10.1146/annurev-psych-120710-100328

Schmiedek, F., Li, S.-C., and Lindenberger, U. (2009). Interference and facilitation in spatial working memory: age-associated differences in lure effects in the n-back paradigm. *Psychol. Aging* 24, 203–210. doi: 10.1037/a0014685

Schurman, J., Brungart, D., and Gordon-Salant, S. (2014). Effects of masker type, sentence context, and listener age on speech recognition performance in 1-back listening tasks. *J. Acoust. Soc. Am.* 136, 3337. doi: 10.1121/1.4901708

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J. Speech Hear. Res.* 28, 455–462. doi: 10.1044/jshr.2803.455

Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., and Gwaltney, C. A. (1999). Monosyllabic word recognition at higher-than-normal speech and noise levels. *J. Acoust. Soc. Am.* 105, 2431–2444.

Tun, P. A., and Wingfield, A. (1999). One voice too many: adult age differences in language processing with different types of distracting sounds. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 54B, P317–P327. doi: 10.1093/geronb/54B.5.P317

Unsworth, N., and Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* 114, 104–132. doi: 10.1037/0033-295X.114.1.104

Verhaeghen, P., and Basak, C. (2005). Ageing and switching of the focus of attention in working memory: results from a modified N-back task. *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* 58A, 134–154. doi: 10.1080/02724980443000241

Vickers, D., and Packer, J. S. (1982). Effects of alternating set for speed or accuracy on response time, accuracy, and confidence in a unidimensional discrimination task. *Acta Psychol.* 50, 179–197.

Wingfield, A., Lindfield, K. C., and Goodglass, H. (2000). Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition. *J. Speech Lang. Hear. Res.* 43, 915–925. doi: 10.1044/jslhr.4304.915

Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *J. Neurosci.* 32, 14010–14021. doi: 10.1523/JNEUROSCI.1528-12.2012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kidd and Humes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 05 June 2015
doi: 10.3389/fpsyg.2015.00721

The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults

Karen A. Doherty[1]* *and Jamie L. Desjardins*[2]

[1] Department of Communication Sciences and Disorders, Syracuse University, Syracuse, NY, USA, [2] Department of Rehabilitation Sciences, University of Texas at El Paso, El Paso, TX, USA

Untreated hearing loss can interfere with an individual's cognitive abilities and intellectual function. Specifically, hearing loss has been shown to negatively impact working memory function, which is important for speech understanding, especially in difficult or noisy listening conditions. The purpose of the present study was to assess the effect of hearing aid use on auditory working memory function in middle-aged and young-older adults with mild to moderate sensorineural hearing loss. Participants completed two objective measures of auditory working memory in aided and unaided listening conditions. An aged matched control group followed the same experimental protocol except they were not fit with hearing aids. All participants' aided scores on the auditory working memory tests were significantly improved while wearing hearing aids. Thus, hearing aids worn during the early stages of an age-related hearing loss can improve a person's performance on auditory working memory tests.

Keywords: age-related hearing loss, presbycusis, aging, hearing aids, working memory

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Jerker Rönnberg,
Linköping University, Sweden
Kathryn Arehart,
University of Colorado, USA

***Correspondence:**
Karen A. Doherty,
Department of Communication Sciences and Disorders, Syracuse University, 621 Skytop Road, Syracuse, NY 13244, USA
kadohert@syr.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *20 February 2015*
Accepted: *14 May 2015*
Published: *05 June 2015*

Citation:
Doherty KA and Desjardins JL (2015) The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults. Front. Psychol. 6:721. doi: 10.3389/fpsyg.2015.00721

Introduction

Age-related hearing loss in middle-aged (MA) and young-older (YO) adults is a public health problem in the U.S. affecting 20% of people between 45–59 years of age and 33% of people in their sixties (National Institute on Deafness and Other Communication Disorders, 2012; Cruickshanks et al., 2003; Nash et al., 2011). Age-related hearing loss is initiated peripherally in the auditory system, and involves hair cell loss, a decline in the cochlear metabolic system, and a loss of spiral ganglion neurons (Frisina and Walton, 2001). The peripheral loss begins in the high frequency regions of the peripheral auditory system and projects to the high frequency regions of the brain, which can induce reorganization of auditory cortical frequency maps (Robertson and Irvine, 1989; Harrison et al., 1991). Due to its gradual onset, mild age-related hearing loss often goes unnoticed. Although signs for early hearing loss exist, many people are unaware of them or choose not to acknowledge them. Instead, they will place the onus of their communication problems on others. For example, individuals with hearing impairment will often suggest people mumble, do not speak clearly, or speak too softly.

Currently, hearing aids are the primary treatment for an age-related hearing loss. However, the uptake rates for adult hearing aid use are low; 20% for all hearing impaired adults, and 15% for adults with hearing loss in their fifties (Lin, 2011). Furthermore, on average, it takes individuals about 10 years from the time they become aware of their hearing problems to when they seek

treatment (Davis et al., 2007). This is concerning because age-related hearing loss can be a serious communication disorder that when left untreated can negatively impact a person's social, and psychological function (for a review, see National Council on the Aging, 1999). Untreated hearing loss has also been related to cognitive function (Lindenberger and Baltes, 1994a; Pichora-Fuller and Singh, 2006; Arlinger et al., 2009; Lin et al., 2013). For example, Lindenberger and Baltes (1994b) found that peripheral auditory thresholds were significantly related to processing speed, working memory, and reasoning in 156 individuals who were 70 years and older. Similarly, van Boxtel et al. (2000) reported a significant relationship between auditory function and verbal memory performance in 453 individuals between 23 and 82 years of age. Recently, Desjardins and Doherty (2013) found that working memory, processing speed and selective attention abilities were significantly associated with older hearing impaired adults' speech recognition performance in background noise.

It has been suggested that a lack of auditory input from an untreated hearing loss could negatively affect the neural networks involved in certain cognitive abilities (Sekuler and Blake, 1987; Lindenberger and Baltes, 1994a; Belin et al., 1999; Wong et al., 2010). That is, a perceptual decline could result in a permanent cognitive decline (*deprivation hypothesis*, Baltes and Lindenberger, 1997). It has also been suggested that even mild hearing loss could lead to a decline in cognitive performance because the cognitive resources normally used for higher-level comprehension, like storing auditory information into memory, must be used by the individual to accurately decode and perceive the speech signal (Pichora-Fuller and Singh, 2006; Rönnberg et al., 2008, 2013; Tun et al., 2009; Gosselin and Gagne, 2011; Desjardins and Doherty, 2013, 2014).

Fortunately, there is evidence that hearing aid use may improve older adults' performance on auditorily presented cognitive tests because the amplified signal likely improves an individual's perception of instructions and test items (Mulrow et al., 1990; Allen et al., 2003; and Weinstein and Amsel, 1986). Allen et al. (2003) reported reduced rates of decline in cognitive screening scores for dementia over a 6-month period following intervention with hearing aids in a group of older adults. Mulrow et al. (1990) found improved performance with the use of hearing aids on a general cognitive measure in adults in their 70 s with moderate sensorineural hearing loss. In addition, a group of older people with dementia were reclassified to a less severe category of dementia when retested with amplification (Weinstein and Amsel, 1986). Hearing aid use has also been shown to reduce listening effort on a speech recognition in noise listening task (Desjardins and Doherty, 2013, 2014). However, other studies have shown that hearing aid use had no effect on older hearing impaired listeners' performance on visual measures of working memory and executive function (Tesch-Romer, 1997; van Hooren et al., 2005). Thus, studies that have examined the effects of hearing aid use on cognition have yielded different results.

In the present study, we examined the effect of hearing aid use on auditory tests of working memory in MA and YO adults. Despite the interest in the association between hearing impairment and cognitive function, only a few studies have investigated whether use of amplification improves working memory performance (Tesch-Romer, 1997; van Hooren et al., 2005). In addition, much of what we know about the negative effects of untreated hearing loss and the potential benefit of hearing aids to offset these effects is based on studies that include participants with an average age of 70 years leaving the earlier stages of age-related hearing loss less understood. We specifically chose to assess working memory function in this study because it has been shown to be necessary for effective speech-communication in noise (Baddeley and Hitch, 1974; Gatehouse et al., 2003; Humes et al., 2006; Akeroyd, 2008), and to decline with increasing age (Salthouse and Lichty, 1985).

Briefly, working memory is a system for the temporary storage, management, and manipulation of information required for carrying out complex cognitive tasks such as language comprehension (Daneman and Carpenter, 1980). Models of working memory assume that when the capacity limits of working memory are exceeded due to processing demands (e.g., background noise), either comprehension will become slowed or errors will occur (Rabbitt, 1990; Rönnberg, 2003). Thus, an impoverished perceptual input due to background noise or hearing impairment could compromise cognitive performance (Rabbitt, 1968; Lindenberger and Baltes, 1994a; Pichora-Fuller, 2008; Rönnberg et al., 2008, 2013). According to the ease of language understanding model (ELU; Rönnberg et al., 2013), in effort demanding listening situations (e.g., listening to speech in background noise), an individual with a high working memory capacity will be better able to compensate for a distorted signal without exhausting their working memory capacity (i.e., making listening less effortful), compared to an individual with a smaller working memory capacity (Rudner et al., 2011; Ng et al., 2013; Mishra et al., 2014; Rudner and Lunner, 2014). Thus, hearing aids may lessen the cognitive processing resources a hearing impaired listener must expend to understand speech by effectively compensating for an auditory impairment (Desjardins and Doherty, 2014).

Materials and Methods

Participants

There were 24 participants divided among 11 MA adults 50–60 years of age [Mean (M) = 56.6 years, Standard Deviation (SD) = 3.4 years], and 13 YO adults 63–74 years of age (M = 68.7 years, SD = 4.1 years). All of the participants in the current study were part of a larger longitudinal hearing aid study. All participants had at least a mild sensorineural hearing loss, bilaterally (i.e., two out of three thresholds were $>$ 26 dB at 2 kHz, $>$ 30 dB at 3 kHz and/or $>$ 35 dB at 4 kHz), and no more than a 15 dB difference in hearing thresholds between ears at any audiometric frequency. This hearing loss criterion was selected so that the participants' thresholds would be at least $>$ 0.5 standard deviations from the normal hearing thresholds reported for these ages in the Cruickshanks et al. (2003) study. Mean pure-tone thresholds for the MA and YO participants averaged across the left and right ears are shown in **Figure 1**.

Two age-matched control groups were also included in this study. The purpose of the control participants was to ensure that any significant changes measured in the experimental groups

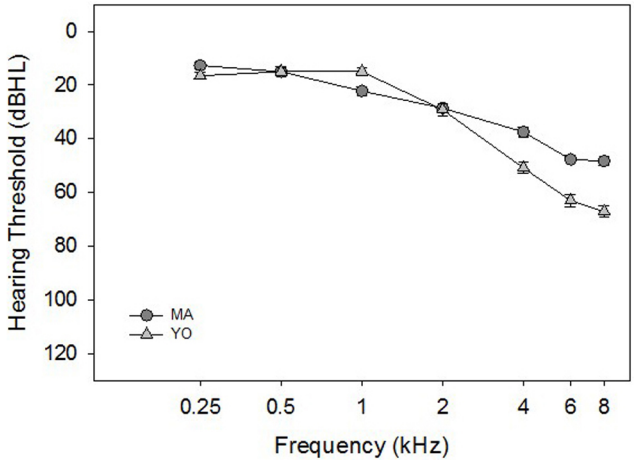


FIGURE 1 | Mean pure-tone thresholds (in dB HL) averaged across the right and left ears for the MA (circles), and YO (triangles) participants. Error bars represent ± 1 SD.

over the 6 week study period were not a result of normal test re-test variability on the experimental test measures. The control groups consisted of a group of 8 MA (C-MA) (Mean Age = 55 years, SD = 2.9 years) adults, and a group of 8 YO (C-YO) (Mean Age = 67 years, SD = 3.1 years) adults. The control participants were recruited in the same manner as the experimental participants in this study. If a participant did not meet the hearing threshold criteria for being fit with hearing aids, they were assigned to one of the control groups depending upon their age.

None of the participants had worn or tried a hearing aid prior to participating in this study. All participants were native speakers of English and were paid an hourly wage for their participation. Institutional Review Board approval was obtained prior to commencement of this study in accordance with the Syracuse University IRB committee.

Amplification

Middle-aged and YO participants were fitted with ReSound Alera 9 (GN ReSound, Ballerup Denmark) receiver-in-the-canal hearing aids coupled to open dome ear molds, bilaterally. Hearing aid gain was determined based on the Desired Sensation Level (DSL v. 5) prescriptive method (Scollie et al., 2005). DSL targets were generated using Avanti 3.2 software in NOAH 3, and verified with the Audio scan Verifit VF-1 real ear system (Dorchester, ON, Canada). The frequency responses of the hearing aids were adjusted so that the real-ear aided response was within 5 dB across the prescribed target values for 0.25, 0.5, 1, 2 kHz, and within 10 dB for 4 kHz and 6 kHz at an input signal of 70 dB SPL. The hearing aids were set to have two programs: (1) Omnidirectional, (2) Adaptive noise reduction. All other programs and the volume control were disabled. Participants were instructed on the use and care of their hearing aids, and asked to wear the hearing aids for at least 8 h per day, every day, for 6 weeks.

The data-logging feature in the hearing aids was used to track the overall hours of hearing aid use over the 6 week hearing aid trial. The Practical Hearing Aid Skills Test Revised (PHAST-R; Desjardins and Doherty, 2009; Doherty and Desjardins, 2012), an eight item objective assessment that measures basic hearing aid use and care skills, was administered to participants at their initial hearing aid fitting session, after 2 weeks of hearing aid use, and at 6 weeks of hearing aid use. The PHAST-R provided an objective measure of the participant's ability to correctly use and care for their hearing aids. After each administration of the PHAST-R, participants were reinstructed on tasks they did not perform correctly or know how to perform.

Test Measures

Working memory function was measured using an auditory version of The Reading Span Test (Daneman and Carpenter, 1980; Pichora-Fuller et al., 1995), and an auditory version of the *n*-back task (*N*-backer; Monk et al., 2011).

Listening Span Test

The Listening Span Test, which is an auditory version of the Reading Span Test, was selected to measure working memory because the Reading Span Test has been shown to be one of the best predictors of speech recognition performance in noise in hearing impaired adults (Akeroyd, 2008; Rönnberg et al., 2010; Desjardins and Doherty, 2013; Ng et al., 2013). The methods used to administer the Listening Span Test in the current study have methodological similarities to those reported for the auditory reading span test in previous studies (Pichora-Fuller et al., 1995; Sarampalis et al., 2009; Ng et al., 2013, 2015). The Listening Span Test in the present study consists of sentences from the revised Speech Perception in Noise (R-SPIN) test (Bilger et al., 1984) which is comprised of eight lists of 50 sentences (400 total sentences). Each list of sentences contains 25 high context sentences such that the final-word in the sentence is predictable (e.g., A chimpanzee is an ape) and 25 low context sentences where the final-word is not predictable (e.g., She might have discussed the ape). R-SPIN sentences were recorded by a female talker and digitized using the Computerized Speech Lab (Kay Elemetrics, Montvale, NJ, USA) at a 44,100 Hz sampling rate. They were presented at 70 dBSPL in quiet, and in a speech shaped noise (SSN) at +8 dB signal-to-noise-ratio (SNR). The SSN was generated in MATLAB using a 16 bit, 44. 1 kHz sampling rate, by passing a Gaussian noise through a Finite Impulse Response filter with a magnitude response equal to the Long Term Average Speech Spectrum of the 400 R-SPIN sentences. The +8 dB SNR level was chosen to avoid ceiling and floor effects on speech recognition performance based on pilot data we collected with MA and YO hearing impaired adults.

The R-SPIN sentences were presented to participants in a double walled sound attenuating booth in quiet and in the SSN in a randomized order via a Sony multi-disc CD changer (Sony electronics Inc., Tokyo, Japan) routed through a GSI-61 audiometer to a GSI loudspeaker (Grason-Stadler, Eden Prairie, MN, USA) located 1 meter, at ear level, in front of the participant (0°azimuth). In the SSN condition, the background masker was played continuously throughout the task. Participants were required to repeat the entire R-SPIN sentence they heard during a 4 s interval that followed the presentation of each sentence,

and to remember the final word in each sentence for later recall. The examiner recorded only the final key word in the sentence. The memory task was manipulated by varying the number of sentences in the set (i.e., 2, 4, and 6). After all the sentences in a given set were presented, the experimenter prompted the participant to recall as many of the previously reported final key words as they could, verbally, and in any order. Twenty-four sentences were presented in each of the six experimental conditions (Quiet: set size 2, 4, 6, and Noise: set size 2, 4, 6). Performance on the Reading Span test was computed based on the percent of correctly recalled final key words.

N-back Test

Participants were administered an auditory version of the *n*-back task (*N*-backer; Monk et al., 2011). The *n*-back is a continuous performance task that is commonly used as an assessment in cognitive neuroscience to measure the executive component of working memory (for reviews, see Kane et al., 2007; Jaeggi et al., 2010). Participants were seated in a double-walled sound attenuating booth and presented a sequence of 25 randomly generated synthesized digits from 1 to 9 using the *N*-backer computer software (Monk et al., 2011) via a computer routed through a GSI-61 audiometer to a GSI loudspeaker (Grason-Stadler, Eden Prairie, MN, USA) located 1 meter, at ear level, in front of the participant (0°azimuth). Each digit was presented with a constant inter-stimulus interval of 2000 ms at 70 dBSPL in quiet and in a SSN at +8 dB SNR in a randomized order. In the SSN condition, the background masker was played continuously throughout the task. The +8 dB SNR level was chosen to avoid ceiling and floor effects on speech recognition performance based on pilot data we collected with this population. Participants were instructed to listen to the stream of randomly presented digits, and to say the digit they heard "1-step" back in time for the 1-back task, and to say the digit they heard "2-steps" back in time for the 2-back task. Participants always completed a practice test session first, during which streams of 10 randomly presented digits were presented in quiet and in noise. Performance on the auditory *n*-back was calculated as the number of correctly recalled digits.

Procedure

Middle-aged and YO participants completed four test sessions over a period of 6 weeks. On the weeks when participants were not seen in the lab, they were contacted via telephone by the examiner to encourage hearing aid use, answer questions, and trouble shoot hearing aid problems. During session 1, all testing was performed unaided, hearing thresholds were obtained at the standard audiometric test frequencies from 0.25 to 8.0 kHz with a GSI-61 audiometer using standard audiometric test procedures (American National Standards Institute [ANSI], 2003). All stimuli were presented at 70 dB SPL, which was above the participants' hearing thresholds. To further ensure that the stimuli were audibile we obtained speech recognition scores for the R-SPIN sentences and the *N*-back digits unaided in quiet and background noise. The Listening Span test and the auditory *N*-back were then administered in quiet and in noise in a randomized order. Session 2 took place within 1 week of session 1. During session 2, the experimental participants were fitted with hearing aids following the hearing aid fitting procedure described in the amplification section. Two weeks after their initial hearing aid fitting, participants returned to the lab to participate in Session 3. During session 3, hearing aid orientation information was reviewed. The PHAST-R (Doherty and Desjardins, 2012) was administered, and participants were reinstructed on the hearing aid use and care skills they did not perform correctly or know how to perform. In addition, participants aided speech recognition in quiet and noise was measured using lists of 24 sentences from the R-SPIN following the standard R-SPIN test instructions (see Bilger et al., 1984). After wearing the hearing aids for 6 weeks, participants returned to the lab for session 4. During session 4, participants were administered the Listening Span Test and the auditory *N*-back while wearing their hearing aids. All testing in background noise was performed with the hearing aids in the adaptive noise reduction setting. At the end of Session 4 participants were asked to return the hearing aids. Participants were then administered the auditory *n*-back test unaided.

The two age-matched control groups followed a similar testing procedure as the two experimental groups, except they were not fitted with amplification. Control participants completed Session 1, as described for the experimental participants. Six weeks after they completed session 1, they returned to the lab for a second test session (i.e., control-session 2). During Control-session 2, control participants were administered the Listening Span Test and the auditory *N*-back in a randomized order.

Results

Speech recognition scores were compared across the control and experimental groups. The mean unaided sentence recognition (R-SPIN) scores were 98% (SD = 4), 95% (SD = 11), 100% (0), and 100% (SD = 0) in quiet and 98% (SD = 3), 94% (SD = 12), 100% (SD = 0), and 100% (SD = 0) in background noise for the MA, YO, C-MA and C-YO groups, respectively. Based on the 95% critical differences for speech recognition percentage scores, there were no significant differences in speech recognition scores among the four groups of participants in this study (Thornton and Raffin, 1978). Mean unaided speech recognition scores for the *N*-back stimuli were 96% (SD = 2.7), 92% (SD = 4), 100% (SD = 0), and 98% (SD = 2) in quiet and 95% (SD = 1.4), 92% (SD = 3.3), 100% (SD = 0), and 96% (SD = 2.3) in background noise for the MA, YO, C-MA, and C-YO groups, respectively. Based on the 95% critical differences for speech recognition percentage scores, there were no significant differences in speech recognition scores among the four groups of participants in this study (Thornton and Raffin, 1978).

On average, MA participants used their hearing aids 12 h per day (SD = 5.5 h), and the YO participants used their hearing aids 11 h per day (SD = 6 h) based on hearing aid data log information. Aided speech recognition scores on the R-SPIN were 100% (SD = 0) in quiet and in background noise for the MA participants, and 100% (SD = 0) and 98% (SD = 1.3) in quiet and in background noise for the YO participants. Mean aided speech recognition scores for the *N*-back digits were 96% (SD = 2.6), and 95% (SD = 1.5) in quiet and 96% (SD = 1.6), 96% (SD = 1.2) in

background noise for the MA and YO groups, respectively. Based on the 95% critical differences for speech recognition percentage scores, there were no significant difference between aided and unaided recognition of R-Spin sentences and *N*-back digits for either group of listeners (Thornton and Raffin, 1978).

Listening Span Test

Working memory function was assessed using the Listening Span test in quiet and in noise with and without hearing aids. Mean scores and standard errors of the mean on the Listening Span test in quiet and in noise, collapsed across context are shown in **Figure 2**. To compare differences in working memory across factors and participant groups, a 3 × 2 × 2 × 2 × 2 full factorial repeated measures analysis of variance (RMANOVA) was performed on the factors span (2 span, 4 span, 6 span), listening condition (quiet, noise), amplification (unaided, aided), context (low and high) and group (MA, YO). Greenhouse-Geisser corrections (Greenhouse and Geisser, 1959) were used to correct sphericity violations throughout the analyses where indicated. All *post hoc* comparisons were completed using the Bonferroni adjustment for multiple comparisons.

There was a significant two-way interaction of Span × Group [$F(2, 20) = 9.6$; $p = 0.001$; partial eta-squared = 0.50]. *Post hoc* analysis indicated significant group differences for the 4 and 6 span conditions but, not for the 2-span condition. YO participants scored significantly lower on the Listening Span test (i.e., poorer working memory performance) in the 4 span ($p = 0.003$) and 6 span ($p = 0.001$) conditions compared to the MA participants in both quiet and noise. There was also a significant two-way interaction between Listening Condition × Amplification [$F(1, 21) = 4.8$; $p = 0.02$; partial eta-squared = 0.20]. MA and YO participants scores on the Listening Span test were significantly ($p < 0.001$) higher (i.e., better working memory performance) with amplification in the 4 and 6 span conditions but, only in the background noise listening condition. Interestingly, while their performance was improved with hearing aids, the YO participants' aided scores on the Listening Span Test approximated the unaided scores of the MA participants in the noisy listening condition.

Two age-matched control groups were used in this study to ensure significant changes in the experimental group were not a result of simply being re-tested on the Listening Span test over the 6 weeks. Mean Listening Span test scores for the MA and YO control participants in quiet and noise at their initial test session and at the second test session, which occurred 6 weeks later, are shown in **Figure 3**. To compare differences in scores on the Listening Span test over time for the MA and YO control groups, a 3 × 2 × 2 × 2 full factorial RMANOVA was performed on the factors span (2 span, 4 span, 6 span), listening condition

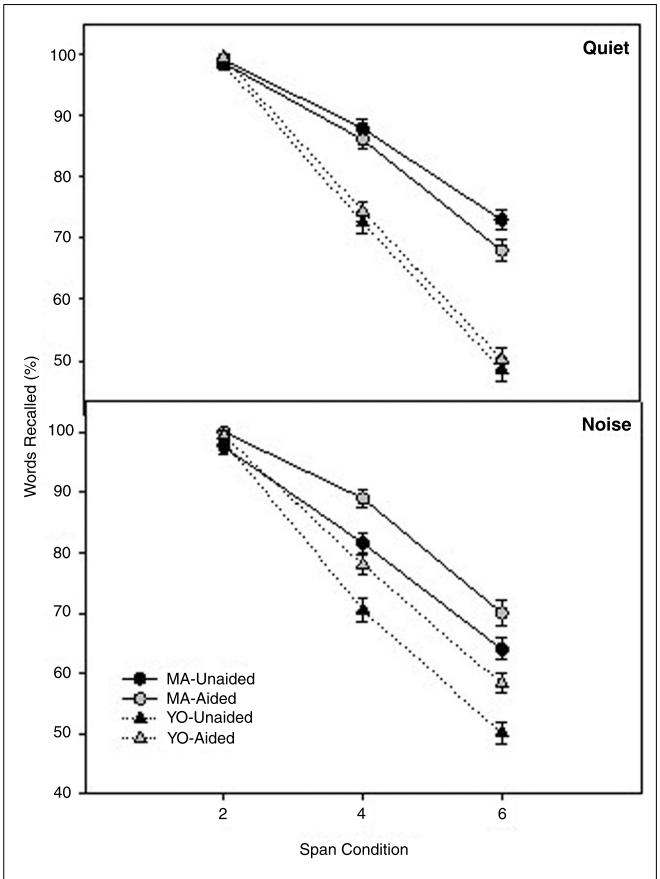


FIGURE 2 | Mean unaided and aided Listening Span test scores for the MA (circles) and YO (triangles) participants in quiet (top) and noise (bottom). Error bars represent ± 1 SE.

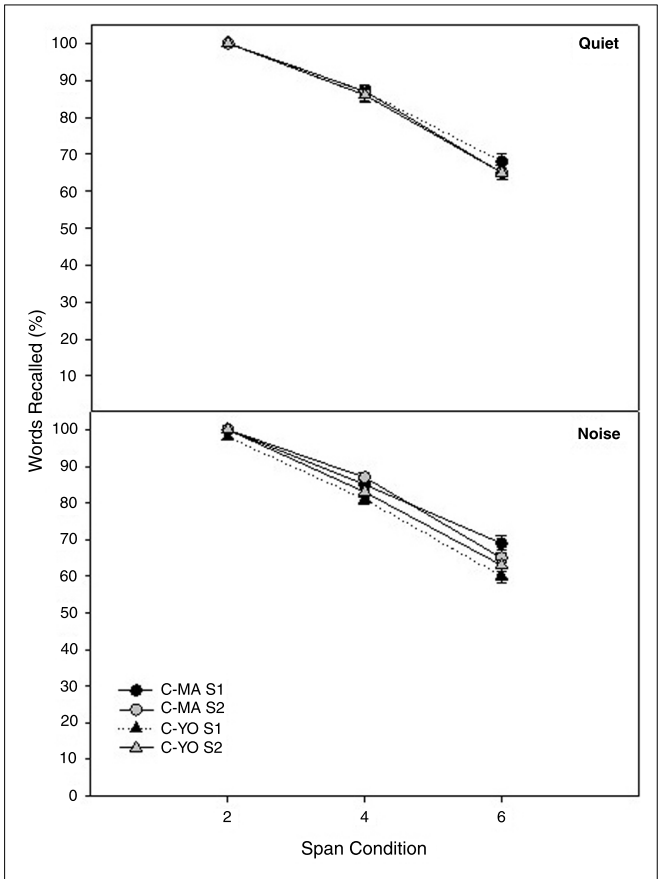


FIGURE 3 | Mean Listening Span test scores in quiet (top) and noise (bottom) for the two age-matched control groups [C-MA (circles) and C-YO (triangles)] for test sessions 1 and 2. Error bars represent ± 1 SE.

(quiet, noise), test session (session 2, session 4), and group (C-MA, C-YO). There was a significant main effect of span [$F(2, 26) = 48.29$; $p < 0.001$; partial eta squared = 0.79]. Both groups of participants scored higher on the 2 span condition than the 4 and 6 span conditions ($p < 0.001$). All other main effects, two-way and three-way interactions were not significant ($P > 0.05$).

N-back Test

Participants' mean unaided and aided scores on the auditory 1-back and 2-back in quiet and in background noise are displayed in **Figure 4**. To compare differences in performance on the auditory *n*-back across factors and participant groups, a 2 × 2 × 2 × 2 full factorial RMANOVA was performed on the factors back (1, 2) listening condition (quiet, noise), amplification (unaided, aided), and group (MA, YO). There was a significant three-way interaction of amplification × back × group [$F(1, 19) = 7.2$; $p = 0.01$; partial eta squared = 0.3]. *Post hoc* analysis indicated that the YO group scored significantly ($p < 0.001$) higher with hearing aids than without hearing aids in both the quiet and noisy listening conditions in the 1-back condition. However, there were no significant ($p > 0.05$) differences in 1-back scores for the MA participants in the quiet or background noise conditions with hearing aid use. Also, no significant ($P > 0.05$) differences were observed between aided and unaided performance in quiet and noise on the 2-back for either MA or YO participants.

In **Figure 5** the mean 1-back and 2-back scores are shown for the MA and YO older control participants in quiet and noise at the initial test session and at a second test session which occurred 6 weeks later. To compare differences in scores on the auditory *n*-back over time for the MA and YO control groups, a 2 × 2 × 2 × 2 full factorial RMANOVA was performed on the factors listening condition (quiet, noise), test session (session 2, session 4), and group (C-MA, C-YO). There were no significant ($p > 0.05$) main effects, two-way or three-way interactions.

We also compared participants' unaided auditory working memory performance on the *N*-back test in quiet and noise pre-fit and post-fit (6 weeks) to determine whether there was a cognitive transfer after using hearing aids for 6 weeks. A RMANOVA was performed on the within subject factors session (pre-fit, post-fit) and listening condition (quiet, noise) and the between subject factor group (MA, YO, C-MA, C-YO). There was no significant interaction between session and group [$F(3, 33) = 1.25$, $p = 0.31$, effect size = 0.1]. Therefore, it appears that the effect of hearing aids on working memory is more perceptual in that the benefit from amplification was directly related to the improved transfer of the signal and not cognitive transfer because wearing the hearing aids for 6 weeks did not change unaided working memory.

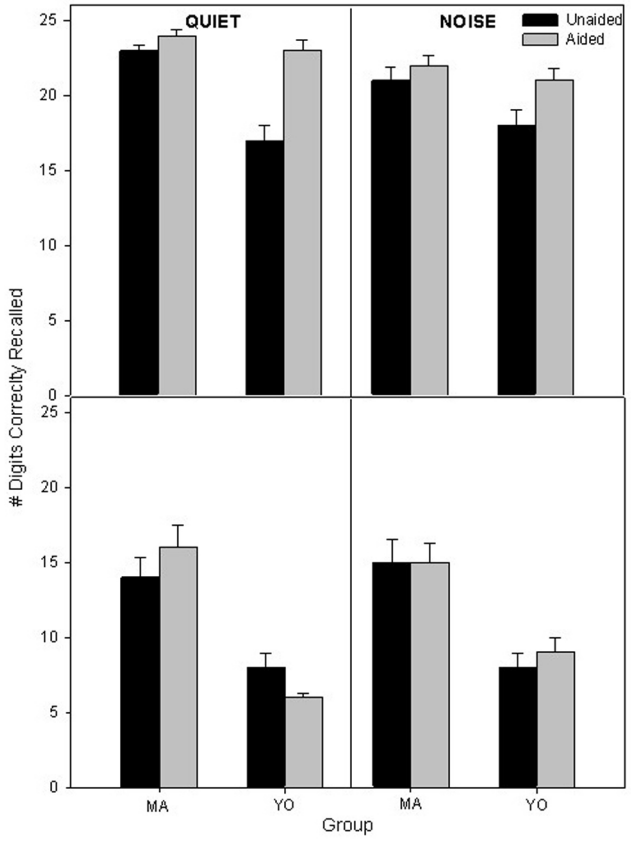


FIGURE 4 | Mean unaided and aided 1-back (top) and 2-back scores (bottom) in quiet and noise for the MA and YO participants. Error bars represent ± 1 SE.

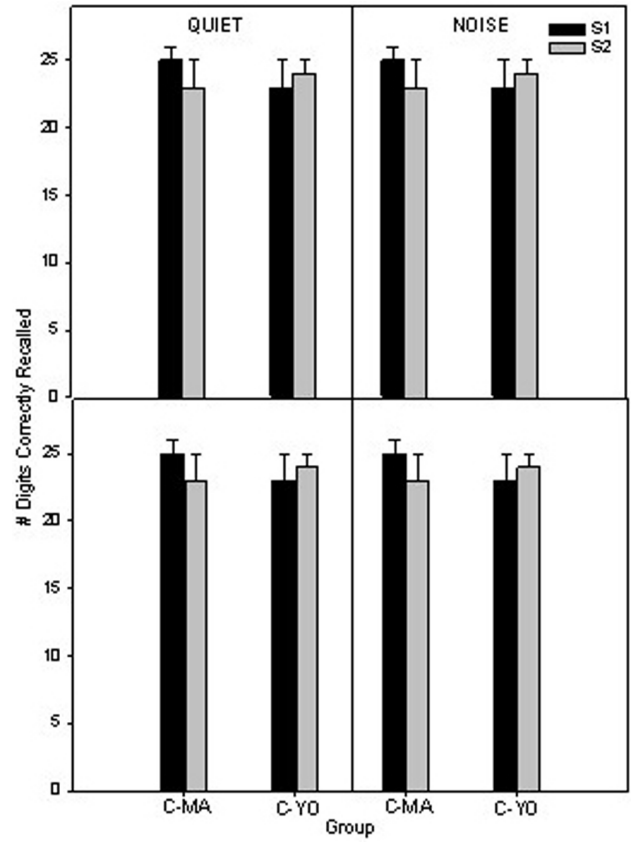


FIGURE 5 | Mean 1-back scores (top) and 2-back scores (bottom) in quiet and noise for the two age-matched control groups (C-MA and C-YO). Error bars represent ± 1 SE.

Discussion

The purpose of this study was to assess the effect of hearing aids on auditory working memory function in MA and YO hearing impaired adults. The main finding of the current study was that MA and YO participants' auditory working memory performance was significantly improved with hearing aid use. This finding is strengthened by the fact that we did not observe any significant changes in working memory performance on the Listening Span test or the auditory *n*-back test in either of the two age-matched control groups who were not fitted with hearing aids. Thus, the significant changes in the experimental group were not a result of simply being re-tested over time.

In this study, we specifically chose to measure the effects of hearing aids on working memory function because, numerous studies have reported on the importance of working memory ability for effective speech-communication in noise (Pichora-Fuller et al., 1995; Gatehouse et al., 2003; Humes et al., 2006; Vaughan et al., 2006; Akeroyd, 2008; Rudner et al., 2011; Besser et al., 2013), and how working memory ability declines with increasing age (Salthouse and Lichty, 1985; Park and Lee, 1999). In a review of twenty studies on speech recognition and cognitive abilities, Akeroyd (2008) found that while hearing sensitivity was the primary predictor of speech recognition performance, working memory capacity, as measured by the reading span test, was the second most important predictor. Long term memory is another factor that could influence listening under suboptimal conditions, e.g., in background noise, or with a hearing loss (Sörqvist and Rönnberg, 2012). Rudner et al. (2011) described how long term memory is used to help infer and construct the meaning of a target message by retrieving phonological, lexical, and semantic representations from an individual's long term memory. However, the current study focused on measuring auditory working memory.

In the current study, both the MA and YO participants' Listening Span test scores were significantly higher with hearing aid use. Although the YO participants working memory performance was improved with hearing aids, their performance never achieved the level to that of the MA group. Interestingly, the YO participants *aided* scores on the reading span test approximated the *unaided* scores of the MA group. This suggests amplification may reduce the confounding effect of hearing loss on apparent early age-related decline in auditory working memory function.

Significant improvements in working memory performance with hearing aids for both MA and YO groups on the auditory Reading Span test were only evident when memory performance was tested in background noise, even though their speech recognition scores were excellent in both quiet and noise. This result largely supports the *effortfulness hypothesis*: the theory that the extra effort that a hearing-impaired listener must expend to successfully understand speech comes at the cost of cognitive processing resources that might otherwise be available for encoding the speech content in memory (Rabbitt, 1968; Kahneman, 1973; Wingfield et al., 2005; Tun et al., 2009). In other words, speech understanding in everyday life is influenced by both bottom-up and top-down cognitive functions that moderate the processing of auditory information (Gatehouse et al., 2003; Humes et al., 2006; Vaughan et al., 2006; Desjardins and Doherty, 2013, 2014). Because speech contains redundant information a hearing-impaired individual can cognitively compensate by "filling in" missed information. Thus, top-down cognitive compensations can effectively mask a peripheral hearing loss and help the hearing impaired listener function more effectively in everyday listening situations (e.g., Tun et al., 2009; Gosselin and Gagne, 2011; Rudner et al., 2011; Desjardins and Doherty, 2013; Rönnberg et al., 2013; Zekveld et al., 2013). For example, in the present study the cognitive demand of the Listening Span Test was greater in noise than in quiet, and therefore required listeners to use more cognitive resources (Murphy et al., 1999; Desjardins and Doherty, 2013, 2014). It is likely that we did not observe significant improvements in working memory scores with hearing aids in the quiet listening condition because the cognitive load did not exceed the individual's cognitive capacity. This result is consistent with a recent study by Mishra et al. (2014) that found residual cognitive capacity, termed Cognitive Spare Capacity, was not reduced in an older group of hearing impaired adults when listening conditions were optimal, but was reduced when the speech was presented in background noise.

Hearing aid use also improved participants' performance on the auditory 1-back task but, no improvements were observed on the 2-back task. On the 1-back only improvements in task scores were observed for the YO group. Interestingly, unlike their performance on the Listening Span Test, the YO participants' aided scores on the 1-back were significantly higher in both quiet and in background noise compared to their unaided scores. It is not too surprising that we observed a different pattern of results on the auditory 1-back compared to the Listening Span test. Several studies which have measured the convergent validity of the *n*-back task with other measures of working memory (see Kane et al., 2007) have largely revealed weak or modest correlations between individuals' performance on the *n*-back task and performance on other standard, accepted assessments of working memory (Kane et al., 2007; Jaeggi et al., 2010). This is because, performance on the *n*-back task seems to be more closely correlated with performance on measures of fluid intelligence than it is with performance on other measures of working memory (Jaeggi et al., 2010). This is interesting because the 1-back, although having a low cognitive load, proved to be a more sensitive measure for assessing hearing aid use in our older group of adults, as significant differences with amplification were observed in both quiet and noisy listening situations. However, it was not sensitive to changes in auditory working memory function with hearing aid use in the MA participants. It is likely that we did not see significant improvements with amplification on the 1-back in the MA group because of a ceiling effect on the task, as their unaided performance was already near excellent. It is somewhat difficult to interpret why there was no improvement with amplification on the 2-back task for either the MA or YO participants. Perhaps the perceptual benefit of amplification could not improve performance on a task with such a high cognitive load.

If the benefit from amplification on the 1-back test was more of a cognitive transfer effect, then the participants' *unaided* working memory performance should have improved after wearing the hearing aids for 6 weeks, which did not occur. Therefore, the effects of amplification was more perceptual (immediate effect on encoding of working memory) than a cognitive transfer (long-term). Another way to measure this would have been to measure aided and unaided performance on the cognitive tests at weeks 1 and 6, and assess if the amount of hearing aid benefit increased over time. However, we did not obtain aided scores on week 1 because this was part of a larger longitudinal hearing aid study, which did not include aided testing at week 1. Regardless, results from the present study indicate that some type of frequency shaping/amplification should be used when testing auditory working memory in hearing-impaired adults, even with a mild degree of hearing loss, to reduce the potential negative effect a degraded peripheral representation of the signal could have on cognitive test scores.

Using hearing aids in the early stages of age-related hearing loss, even when hearing loss is mild, can improve performance on auditory working memory tests in quiet and in background noise. The majority of MA and YO adults are still in the workforce, and although they may be able to "get by" without a hearing aid, it is important to consider the impact of their hearing loss on their working memory function. Although results from this study indicate wearing hearing aids can have a positive impact on working memory performance, future research should investigate if using hearing aids during the earlier stages of age-related hearing loss can reduce or even prevent some of the perceptual changes that result from auditory deprivation (Thai-Van et al., 2010).

Acknowledgments

This project was funded by a P30 Grant # P30AG034464 from the National Institutes of Health/ National Institute of Aging. We would like to thank the GN ReSound company for supplying the hearing aids used in this study.

References

Allen, N. H., Burns, A., Newton, V., Hickson, F., Ramsden, R., Rogers, J., et al. (2003). The effects of improving hearing in dementia. *Age Ageing* 32, 189–193. doi: 10.1093/ageing/32.2.189

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

American National Standards Institute [ANSI]. (2003). *American National Specification of Hearing Aid Characteristics.* New York: Acoustical Society.

Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, K. (2009). The emergence of cognitive hearing science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x

Baltes, P. B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol. Aging* 12, 12.

Baddeley, A. D., and Hitch, G. (1974). Working memory. *Psychol. Learn. Motiv.* 8, 47–89. doi: 10.1016/S0079-7421(08)60452-1

Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., and Pike, B. (1999). Event-related fMRI of the auditory cortex. *Neuroimage* 10, 417–429. doi: 10.1006/nimg.1999.0480

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., and Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *J. Speech Hear. Res.* 27, 32–48. doi: 10.1044/jshr.2701.32

Cruickshanks, K. J., Tweed, T. S., Wiley, T. L., Klein, B. E. K., Klein, R., Chappell, R., et al. (2003). The 5-year incidence and progression of hearing loss: the epidemiology of hearing loss study. *Arch. Otolaryngol. Head Neck Surg.* 129, 1041–1046. doi: 10.1001/archotol.129.10.1041

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Davis, A., Smith, P., Ferguson, M., Stephens, D., and Gianopoulos, I. (2007). Acceptability, benefit, and costs of early screening for hearing disability: a study of potential screening tests, and models. *Health Technol. Assess.* 11, 1–294. doi: 10.3310/hta11420

Desjardins, J. L., and Doherty, K. A. (2009). Do experienced hearing aid users know how to use their hearing aids correctly? *Am. J. Audiol.* 18, 69–76. doi: 10.1044/1059-0889(2009/08-0022)

Desjardins, J. L., and Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear Hear.* 34, 3, 261–272. doi: 10.1097/AUD.0b013e31826d0ba4

Desjardins, J. L., and Doherty, K. A. (2014). The effect of hearing aid noise reduction on listening effort in hearing-impaired adults. *Ear Hear.* 35, 600–610. doi: 10.1097/AUD.0000000000000028

Doherty, K. A., and Desjardins, J. L. (2012). The practical hearing aids skills test—revised. *Am. J. Audiol.* 21, 100–105. doi: 10.1044/1059-0889(2012/11-0019)

Frisina, R. D., and Walton, J. P. (2001). "Aging of the mouse central auditory system," in *Handbook of Mouse Auditory Research: From Behavior to Molecular Biology*, ed. J. P. Willott (Boca Raton: CRC press), 339–379. doi: 10.1201/9781420038736.ch24

Gatehouse, S., Naylor, G., and Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Int. J. Audiol.* 42(Suppl. 1), S77–S85. doi: 10.3109/14992020309074627

Gosselin, P., and Gagne, J. P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *J. Speech Lang. Hear. Res.* 54, 944–958. doi: 10.1044/1092-4388(2010/10-0069)

Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823

Harrison, R. V., Nagasawa, A., Smith, D. W., Stanton, S., and Mount, R. J. (1991). Reorganization of auditory cortex after neonatal high frequency cochlear hearing loss. *Hear. Res.* 54, 11–19. doi: 10.1016/0378-5955(91)90131-R

Humes, L. E., Lee, J. H., and Coughlin, M. P. (2006). Auditory measures of selective and divided attention in young and older adults using single-talker competition. *J. Acoust. Soc. Am.* 120, 2926–2937. doi: 10.1121/1.2354070

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., and Meier, B. (2010). The concurrent validity of the *N*-back task as a working memory measure. *Memory* 18, 394–412. doi: 10.1080/09658211003702171

Kane, M. J., Conway, A. R. A., Miura, T. K., and Colflesh, G. J. H. (2007). Working memory, attention control, and the *N*-back task: a question of construct validity. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 615–622. doi: 10.1037/0278-7393.33.3.615

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Lin, F. (2011). Hearing loss and cognition among older adults in the United States. *J. Gerontol. Ser. A* 66A, 1131–1136. doi: 10.1093/gerona/glr115

Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA Intern. Med.* 173, 293–299. doi: 10.1001/jamainternmed.2013.1868

Lindenberger, U., and Baltes, P. B. (1994a). "Aging and intelligence," in *The Encyclopedia of Human Intelligence*, eds R. J. Sternberg, S. J. Ceci, J. L. Horn, E. Hunt, J. D. Matarazzo, and S. Scarr (New York: Macmillan), 52–66.

Lindenberger, U., and Baltes, P. B. (1994b). Sensory functioning and intelligence in old age: a strong connection. *Psychol. Aging* 9, 339.

Mishra, S., Stenfelt, S., Lunner, R., Rönnberg, J., and Rudner, M. (2014). Cognitive spare capacity in older adults with hearing loss. *Front. Aging Neurosci.* 6:69. doi: 10.3389/fnagi.2014.00096

Monk, A. F., Jackson, D., Nielsen, D., Jefferies, E., and Olivier, P. (2011). *N*-backer: an auditory *n*-back task with automatic scoring of spoken responses. *Behav. Res. Methods* 43, 888–896. doi: 10.3758/s13428-011-0074-z

Mulrow, C. D., Aguilar, C., Endicott, J. E., Tuley, M. R., Velez, R., Charlip, W. S., et al. (1990). Quality-of-life changes and hearing impairment: a randomized trial. *Ann. Intern. Med.* 113, 188–194. doi: 10.7326/0003-4819-113-3-188

Murphy, D. R., McDowd, J. M., and Wilcox, K. A. (1999). Inhibition and aging: similarities between younger and older adults as revealed by the processing of unattended auditory information. *Psychol. Aging* 14, 44–59. doi: 10.1037/0882-7974.14.1.44

Nash, S. D., Cruickshanks, K. J., Klein, R., Klein, B. E. K., Nieto, F. J., Huang, G. H., et al. (2011). The prevalence of hearing impairment and associated risk factors: the beaver dam offspring study. *Arch. Otolaryngol. Head Neck Surg.* 137, 432–439. doi: 10.1001/archoto.2011.15

National Council on the Aging, (1999). *The Impact of Untreated Hearing Loss in Older Americans.*. Washington, DC: Seniors Research Group.

National Institute on Deafness and Other Communication Disorders, (2012). Hearing, Ear Infections, and Deafness. Available at: http://www.nidcd.nih.gov/health/hearing/Pages/Default.aspx [accessed 14 June 2013].

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.* 52, 433–441. doi: 10.3109/14992027.2013.776181

Ng, E. H. N., Rudner, M., Lunner, T., and Rönnberg, J. (2015). Noise reduction improves memory for target language speech in competing native but not foreign language speech. *Ear Hear.* 36, 82–91. doi: 10.1097/AUD.0000000000000080

Park, K., and Lee, S. (1999). "Selective attention for robust speech recognition in noisy environments," *International Joint Conference on Neural Networks (IJCNN'99)*, Washington, DC, 5, 3014–3019.

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *Int. J. Audiol.* 47, S72–S82. doi: 10.1080/14992020802307404

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Rabbitt, P. (1968). Recognition: memory for words currently heard in noise. *Psychon. Sci.* 37, 87–97.

Rabbitt, P. M. A. (1990). Age, IQ and awareness of errors. *Ergonomics* 33, 1291–1305. doi: 10.1080/00140139008925333

Robertson, D., and Irvine, D. R. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.* 282, 456–471. doi: 10.1002/cne.902820311

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42, S68–S76. 456–471. doi: 10.3109/14992020309074626

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi: 10.1080/14992020802301167

Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12, 263–269. doi: 10.4103/1463-1741.70505

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rudner, M., and Lunner, T. (2014). Cognitive spare capacity and speech communication: a narrative overview. *Biomed. Res. Int.* 2014:869726. doi: 10.1155/2014/869726

Rudner, M., Rönnberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4

Salthouse, T. A., and Lichty, W. (1985). Tests of the neural noise hypothesis of age-related cognitive change. *J. Gerontol.* 40, 443–450. doi: 10.1093/geronj/40.4.443

Sarampalis, A., Sridhar, K., and Edwards, B. Hafter, E. (2009). Objective measures of listening effort: effects of background noise and noise reduction. *J. Speech Lang. Hear. Res.* 52, 1230–140. doi: 10.1044/1092-4388(2009/08-0111)

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., et al. (2005). The desired sensation level multistage input/output algorithm. *Trends Amplif.* 9, 159–197. doi: 10.1177/108471380500900403

Sekuler, R., and Blake, R. (1987). Sensory underload. *Psychol. Today* 21, 48–51.

Sörqvist, P., and Rönnberg, J. (2012). Episodic long-term memory of spoken discourse masked by speech: what is the role for working memory capacity? *J. Speech Lang. Hear. Res.* 55, 210–218.

Tesch-Romer, C. (1997). Psychological effects of hearing aid use in older adults. *J. Gerontol. Psychol.* 52, 127–138.

Tun, P. A., McCoy, S., and Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychol. Aging* 24, 761–766. doi: 10.1037/a0014802

Thai-Van, H., Veuillet, E., Norena, A., Guiraud, J., and Collet, L. (2010). Plasticity of tonotopic maps in humans: influence of hearing loss, hearing aids and cochlear implants. *Acta Otolaryngol.* 130, 333–337. doi: 10.3109/00016480903258024

Thornton, A. R., and Raffin, M. J. (1978). Speech-discrimination scores modeled as a binomial variable. *J. of Speech Lang. Hear. Res.* 2, 507–518.

van Hooren, S., Anteunis, L., Valentijn, S., and Bosma, H. (2005). Does cognitive function in older adults with hearing impairment improve by hearing aid use? *Int. J. Audiol.* 44, 265–271. doi: 10.1080/14992020500060370

van Boxtel, M. P., van Beijsterveldt, C. E., Houx, P. J., Anteunis, L. J., Metsemakers, J. F., and Jolles, J. (2000). Mild hearing impairment can reduce verbal memory performance in a healthy adult population. *J. Clin. Exp. Neuropsychol.* 22, 147–154. doi: 10.1076/1380-3395(200002)22:1;1-8;FT147

Vaughan, N., Storzbach, D., and Furukawa, I. (2006). Sequencing versus nonsequencing working memory in understanding of rapid speech by older listeners. *J. Am. Acad. Audiol.* 17, 506–518. doi: 10.3766/jaaa.17.7.6

Weinstein, B., and Amsel, L. (1986). The relationship between dementia and hearing impairment in the institutionalized elderly. *Cin. Gerontol.* 4, 3–15. doi: 10.1300/J018v04n03_02

Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Wong, P., Ettlinger, M., Sheppard, J., Gunasekera, G., and Dhar, S. (2010). Neuroanatomical characteristics and speech perception in noise in older adults. *Ear Hear.* 31, 471–479. doi: 10.1097/AUD.0b013e3181d709c2

Zekveld, A. A., Rudner, M., Johnsrude, I. S., and Rönnberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *J. Acoust. Soc. Am.* 134, 2225–2234. doi: 10.1121/1.4817926

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Doherty and Desjardins. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 02 June 2014
doi: 10.3389/fpsyg.2014.00531

Gated auditory speech perception: effects of listening conditions and cognitive capacity

Shahram Moradi[1]*, Björn Lidestam[2], Amin Saremi[3,4] and Jerker Rönnberg[1]

[1] Linnaeus Centre HEAD, The Swedish Institute for Disability Research, Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden
[2] Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden
[3] Division of Technical Audiology, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden
[4] Cluster of Excellence "Hearing4all", Department for Neuroscience, Computational Neuroscience Group, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

Edited by:
Mari Tervaniemi, University of Helsinki, Finland

Reviewed by:
Mireille Besson, Institut de Neurosciences Cognitives de la Meditarranée, France
Oded Ghitza, Boston University, USA

***Correspondence:**
Shahram Moradi, Linnaeus Centre HEAD, Department of Behavioral Sciences and Learning, Linköping University, SE-581 83 Linköping, Sweden
e-mail: shahram.moradi@liu.se

This study aimed to measure the initial portion of signal required for the correct identification of auditory speech stimuli (or isolation points, IPs) in silence and noise, and to investigate the relationships between auditory and cognitive functions in silence and noise. Twenty-one university students were presented with auditory stimuli in a gating paradigm for the identification of consonants, words, and final words in highly predictable and low predictable sentences. The Hearing in Noise Test (HINT), the reading span test, and the Paced Auditory Serial Attention Test were also administered to measure speech-in-noise ability, working memory and attentional capacities of the participants, respectively. The results showed that noise delayed the identification of consonants, words, and final words in highly predictable and low predictable sentences. HINT performance correlated with working memory and attentional capacities. In the noise condition, there were correlations between HINT performance, cognitive task performance, and the IPs of consonants and words. In the silent condition, there were no correlations between auditory and cognitive tasks. In conclusion, a combination of hearing-in-noise ability, working memory capacity, and attention capacity is needed for the early identification of consonants and words in noise.

Keywords: gating paradigm, auditory perception, consonant, word, final word in sentences, silence, noise

INTRODUCTION

Previous studies have attempted to establish isolation points (IPs), that is, the initial portion of a specific acoustic signal required for the correct identification of that signal, in silent conditions (see Grosjean, 1980). An *IP* refers to a given point in the total duration of a speech signal (i.e., a word) that listeners are able to correctly guess the identity of that signal with no change in their decision after hearing the reminder of that signal after that given point. In the present study, we investigated the IPs of different types of spoken stimuli (consonants, words, and final words in sentences) in both silence and noise conditions, in order to estimate the extent to which noise delays identification. In addition, a cognitive hearing science perspective was used to evaluate the relationships between explicit cognitive variables (working memory and attentional capacities), speech-in-noise perceptual ability, and IPs of spoken stimuli in both silence and noise.

THE INITIAL PORTION OF STIMULI REQUIRED FOR CORRECT IDENTIFICATION OF CONSONANTS, WORDS, AND FINAL WORDS IN SENTENCES

CONSONANT IDENTIFICATION

The specific combined features of place (the place in the vocal tract that an obstruction occurs), manner (the configuration of articulators, i.e., tongue or lips, when producing a sound), and voicing (absence or presence of vocal fold vibration) constitute a given consonant. Listeners can correctly identify a consonant when these particular features are available (Sawusch, 1977). Smits (2000) reported that the location and spread of features for stops, fricatives, and nasals are highly variable. In a French gating-paradigm study, Troille et al. (2007) showed that for a 120-ms /z/ consonant, identification occurred about 92 ms before its end.

Noise in combination with the acoustic features of consonants may cause a perceptual change, such that the noise may be morphed together with the consonant, masking or adding consonant features, thereby changing the percept into another consonant (Miller and Nicely, 1955; Wang and Bilger, 1973; Phatak and Allen, 2007). As a result, the number of correctly identified consonants in noise is reduced (Wang and Bilger, 1973; Phatak and Allen, 2007). Phatak and Allen (2007) reported that consonant identification in white noise falls into three categories: a set of consonants that are easily confused with each other (e.g., /f v b m/), a set of consonants that are intermittently confused with each other (e.g., /n p g k d/), and a set of consonants that are hardly ever confused with each other (e.g., /t s z /). Based on the results of Phatak and Allen (2007) showing that noise impacts differently on different consonants, one may predict that the influence of noise should be larger for the consonants that are more easily confused with each other. Furthermore, the signal-to-noise ratio (SNR) required for the identification of consonants varies across consonants (Miller and Nicely, 1955; Woods et al., 2010). We therefore expect that, compared with silence, noise will generally delay the correct identification of consonants.

IDENTIFICATION OF ISOLATED WORDS

Word identification requires an association between an acoustic signal and a lexical item in long-term memory (Lively et al., 1994). According to the cohort model (Marslen-Wilson, 1987), initial parts of a speech signal activate several words in the lexicon. As successively more of the acoustic signal is perceived, words in the lexicon are successively eliminated. Word identification occurs when only one word candidate is left to match the acoustic signal. Gating paradigm studies have generally demonstrated that word identification occurs after a little more than half of the duration of the whole word (Grosjean, 1980; Salasoo and Pisoni, 1985).

Identification of isolated words is poorer in noise than in silence (Chermak and Dengerink, 1981; Dubno et al., 2005). As the main constituents of words, some vowels (Cutler et al., 2005) and consonants (Woods et al., 2010) are highly affected by noise. For instance, Parikh and Loizou (2005) showed that whereas /o/ had the lowest identification score in a noisy condition compared to other vowels, /i/ had the highest identification score. Presentation of /o/ in a noisy condition activated perception of other vowels like /U/. Based on the findings of Parikh and Loizou (2005), noise has differential effects on identification of different vowels (similar to consonants), meaning that the combination of vowels and consonants with noise activates other vowels and consonants, which disturbs the mapping of the input signal with the representations in the mental lexicon. We expect that the addition of these noise-induced extra-activated candidates will delay IPs, as more acoustic information will be needed to map the signal with the phonological representations in the mental lexicon. In addition, noise is likely to be detrimental to the success of this mapping, as it results in a lower intelligibility.

IDENTIFICATION OF FINAL WORDS IN SENTENCES

When words are presented in sentences, listeners can benefit from the syntactic structure (Miller and Isard, 1963) and semantic context in congruent sentences (Kalikow et al., 1977), which in turn can speed up target word identification in comparison with word-alone presentation (Miller et al., 1951; Grosjean, 1980; Salasoo and Pisoni, 1985). This improvement in word identification occurs because contextual factors inhibit the activation of other lexical candidates that are a poorer fit for the linguistic context (Marslen-Wilson, 1987).

The predictability of sentences is a key variable for final word identification in sentences. The estimation of word predictability is derived from a "cloze task procedure" (Taylor, 1953) when subjects are asked to perform a sentence completion task with the final word is missing. For instance, the word "bird" in the sentence "a pigeon is a kind of bird" is an example of a highly predictable word but in the sentence "she pointed at the bird" it is as an example of a low predictable word. It should be noted that the highly predictable and low predictable words differ from anomalous words, wherein words are randomly substituted. Regarding the example above, the word "bird" is incongruous in the sentence "The book is a bird." Final words are easier to identify in meaningful sentences than in semantically anomalous sentences (Miller and Isard, 1963). Highly predictable sentence contexts enhance one's capability to disambiguate final words compared with low predictable sentence contexts (Kalikow et al., 1977).

Prior context facilitates word identification in noise (e.g., Grant and Seitz, 2000); when highly predictable sentences are heard, the auditory thresholds for word identification are lowered (Sheldon et al., 2008; Benichov et al., 2012). Final word identification in noise is different from tests on sentence comprehension in noise (e.g., the Hearing in Noise Test [HINT], Nilsson et al., 1994; Hällgren et al., 2006). The latter requires the listener to repeat the entirety of sentences, in an adaptive procedure. However, final word identification tasks are usually presented at a constant SNR, and require participants to predict which word will come at the end of the sentence, and therefore demands less cognitive effort. They thus differ in the retrieval demands they put on explicit resources such as working memory (Rönnberg et al., 2013).

COGNITIVE DEMANDS OF SPEECH PERCEPTION IN SILENCE AND NOISE

According to the Ease of Language Understanding (ELU) model (Rönnberg et al., 2008), working memory acts as an interface between incoming signals and the mental lexicon. Working memory enables the storage and processing of information during online language understanding. In this model, the incoming signal automatically feeds forward at a sub-lexical (syllable) level in rapid succession to match the corresponding phonological representation in the mental lexicon (cf. Poeppel et al., 2008; Rönnberg et al., 2013). This process of syllabic matching is assumed to demand less working memory capacity for normal-hearing people under optimum listening conditions, resulting in rapid and implicit online language processing. However, if the incoming signal is poorly specified or distorted (e.g., in noisy conditions), a mismatch (or non-match, cf. Rönnberg et al., 2013 for a detailed discussion on the match/mismatch issue) will occur with the phonological representation in the mental lexicon. The rapid and implicit process of lexical access is temporarily disturbed under such conditions. In such cases, explicit and deliberate cognitive processes (i.e., inference-making and attentional processing) are invoked to compensate for this mismatch in order to detect or reconstruct the degraded auditory signal. Previous studies have shown that attentional and inference-making processes greatly depend on working memory capacity (Kane and Engle, 2000; De Neys et al., 2003). Independent support for the ELU model (Rönnberg et al., 2008) comes from studies showing two auditory cortical mechanisms of processing: an automatic segregation of sounds, and an attention-demanding network that analyzes the acoustic features of incoming auditory signals (Petkov et al., 2004; Snyder et al., 2006, see also Rönnberg et al., 2013). Röer et al. (2011) reported that auditory distraction disturbs the automatic connection of auditory stimuli to the phonological representations in long-term memory.

Previous research has supported the notion that working memory capacity is crucial for speech perception in adverse listening conditions (for recent reviews, see Rönnberg et al., 2010, 2013; Mattys et al., 2012). Unfavorable listening conditions place higher demands on working memory processing (Lunner et al., 2009), and less resources are therefore available for the storage of incoming signals (Rabbitt, 1968).

Attentional capacity of listeners is also a cognitive function that plays a critical role in speech perception under degraded listening

conditions (Carlyon et al., 2001; Shinn-Cunningham and Best, 2008; Mesgarani and Chang, 2012). In degraded listening conditions, attention is focused on the signal's frequency (Dai et al., 1991), the spatial spectrum (Mondor et al., 1998; Boehnke and Phillips, 1999), one channel of information (Conway et al., 2001), or the switching between channels of information (Colflesh and Conway, 2007). This focus of attention enables the segregation of different types of auditory competitors for speech understanding and subsequent memory encoding (cf. Rönnberg et al., 2008, 2013; Sörqvist and Rönnberg, 2012; Sörqvist et al., 2012).

THE PRESENT STUDY

The general purpose was to study how large the initial portion of the stimulus needs to be in order for correct identification, and therefore how demanding the perception is, as an effect of how easy the signal is to discriminate and predict. IPs refer to how large the initial portion of the entire signal that is needed for correct identification. Hence, IPs specify how much of the entire signal is required for correct identification, and thereby how quickly the stimuli are identified. It can be assumed that the identification of stimuli is less demanding if the stimuli are identified earlier. Therefore, IPs should allow us to estimate the amount of cognitive demand needed for correct identification of speech stimuli in silence versus in noise, which lowers discriminability, and under different levels of predictability (e.g., due to lexical and sentential context). In turn, this should be reflected in correlations with measures of explicit cognitive functions.

The general purpose encompasses two aims. The first aim was to compare the IPs of different types of spoken stimuli (consonants, words, and final words in sentences) in both silence and noise conditions, using a gating paradigm (Grosjean, 1980). Subordinate to this aim were two more specific research questions. Firstly, *how much does noise generally affect IPs*? It was assumed that masking speech with noise would generally delay IPs. Secondly, *how does noise affect IPs when considering linguistic* (i.e., *lexical and sentential*) *context*? In consonant identification, compensatory lexical and contextual resources were not available in the present study. Therefore, listeners had to identify the consonants based on critical cues of their acoustic properties, distributed across their entire durations. In word identification, the masking of consonants and vowels with noise is likely to diminish one's ability to identify the words, or to misdirect the listener to interpret them as other words. However, lexical knowledge may aid listeners (Davis and Johnsrude, 2007), although noise is likely to delay IPs for words (as well as for consonants). In final word identification in sentences, we therefore assumed that the contextual and semantic information inherent in naturalistic sentences would speed up the identification of target words, even in noise, compared to words presented in isolation. Words positioned at the end of sentences that had either a low predictable or a high predictable semantic context were also compared, so as to further test the benefit of contextual support.

The second aim was to investigate the relationship between explicit cognitive functions (capacities of working memory and attention) and the IPs of different types of spoken stimuli (consonants, words, and final words in sentences) in both silence and noise conditions. On the basis of the ELU model (e.g., Rönnberg et al., 2008, 2013) as well as several independent empirical studies (e.g., Petkov et al., 2004; Snyder et al., 2006; Foo et al., 2007; Rudner et al., 2009, 2011), we predicted that significant correlations would exist between performance in tests of attention and working memory and IPs of gated stimuli in noise, but to a relatively lesser extent in silence.

METHODS

Participants

Twenty-one university students (12 males and 9 females) at Linköping University, Sweden were paid to participate in this study. Their ages ranged from 20 to 33 years ($M = 24.6$ years). All of the students were Swedish native speakers that spoke Swedish at home and at the university. According to the Swedish educational system, the students (or pupils) learn English and at least one another language (e.g., German, French, Spanish) in school. The participants reported having normal hearing, normal vision (or corrected-to-normal vision), and no psychological or neurological pathologies. The participants gave consent, pursuant to the ethical principles of the Swedish Research Council (Etikregler för humanistisk-samhällsvetenskaplig forskning, n.d.), the Regional Ethics Board in Linköping, and Swedish practice for research on normal populations.

MEASURES

Gating speech tasks

Consonants. The study employed 18 Swedish consonants presented in vowel-consonant-vowel syllable format (/aba, ada, afa, aga, aja, aha, aka, ala, ama, ana, aŋa, apa, ara, aʈa, asa, aʃa, ata, ava/). The gate size for consonants was set at 16.67 ms. The gating started after the first vowel /a/ and right at the beginning of the consonant onset. Hence, the first gate included the vowel /a/ plus the initial 16.67[1] ms of the consonant, the second gate gave an additional 16.67 ms of the consonant (a total of 33.34 ms of the consonant), and so on. The minimum, average, and maximum total duration of consonants were 85, 198, and 410 ms, respectively. The maximum number of gates required for identification was 25. The consonant gating task took between 40 and 50 min to complete.

Words. The words in this study were chosen from a pool of Swedish monosyllabic words in a consonant-vowel-consonant format that had average to high frequencies according to the Swedish language corpus PAROLE (2011). Forty-six of these words (all nouns) were chosen and divided into two lists (A and B) comprising 23 words each. Both lists were matched in terms of onset phonemes and neighborhood size (i.e., lexical candidates that shared similar features with the target word). Each word used in the present study had a small to average numbers of neighbors (3–6 alternative words with the same pronunciation of the

[1]The rationale for setting gate size to 16.67 ms came from audiovisual gating tasks (See Moradi et al., 2013), to get the same gate size for both conditions (i.e., audiovisual and auditory modalities). By using 120 frames/s for recording visual speech stimuli, 8.33 ms of a visual stimulus is available in each frame (1000 ms/120 frame/s = 8.33 ms). Multiplying 8.33 by 2 (frames), there is 16.67 ms (Please see Lidestam, 2014, for detailed information).

first two phonemes, e.g., the target word /dop/ had the neighbors /dog, dok, don, dos/). For each participant, we presented one list in silence and the other in noise. The presentation of words was randomized across participants. Participants in the pilot studies complained that word identification with the gate size used for consonants (16.67 ms) led to fatigue and a loss of motivation. Therefore, a doubled gate size of 33.3 ms was used for word identification and also we presented the first phoneme (consonant) of each word as a whole, and gating was started from the onset of the second phoneme (vowel) in order to prevent any exhaustion for the participants. The minimum, average, and maximum duration of words were 548, 723, and 902 ms, respectively. The maximum number of gates required for identification was 21. The word gating task took between 35 and 40 min to complete.

Final Words in Sentences. There were two types of sentences in this study, which differed according to how predictable the last word in each sentence was: sentences with a highly predictable (HP) last word (e.g., "Lisa gick till biblioteket för att låna en *bok*"; "Lisa went to the library to borrow a *book*") and sentences with a low predictable (LP) last word (e.g., "I förorten finns en fantastisk *dal*"; "In the suburb there is a fantastic *valley*"). The last (target) word in each sentence was always a monosyllabic noun.

To begin with, we constructed a battery of sentences that had differing predictability levels. This was followed by three consecutive pilot studies for the development of HP and LP sentences. First, the preliminary versions of sentences were presented in written form to some of the staff members at Linköping University in order to grade the predictability level of the target words in each sentence, from 0 (unpredictable) to 10 (highly predictable), and to obtain feedback on the content of the sentences in order to refine them. The sentences with scores over 7 were used as HP sentences, and those with scores below 3 were used as LP sentences. The rational for criterion below 3 for final words in LP sentences was based on our interest to have a minimum predictability in the sentences in order to separate identification of final words in LP sentences from identification of final words in anomalous sentences or identification of isolated-words. We then revised the sentences on the basis of the feedback. A second pilot study was conducted on 15 students at Linköping University to grade the predictability level of the revised sentences in the same way (from 0 to 10). Once again, the sentences with scores over 7 were used as HP sentences, and those with scores below 3 were used as LP sentences. In a third pilot study, the remaining sentences were presented to another 15 students to grade their predictability level. Again, we chose the sentences with scores over 7 as HP sentences, and the sentences with scores below 3 as LP sentences.

In total, there were 44 sentences (22 HP sentences and 22 LP sentences, based on the last word in each sentence). The gating started from the onset of the first phoneme of the target word. Because of the supportive effects of context on word identification, and based on the pilot data, we set the gate size at 16.67 ms to optimize time resolution. The average duration of each sentence was 3030 ms. The minimum, average, and maximum duration for target words at the end of sentences were 547, 710, and 896 ms, respectively. The maximum number of gates required for identification was 54. The gating final-word in sentence task took between 25 and 30 min to complete.

Hearing in Noise Test

We used a Swedish version of the HINT (Hällgren et al., 2006), adapted from Nilsson et al. (1994), to measure the hearing-in-noise ability of the participants. The HINT sentences consisted of three-to-seven word everyday sentences with fluctuating ± 2 dB SNR. The sentences were normalized for naturalness, difficulty, and reliability. The sentences were read aloud by a female speaker. In the present study, we used one list consisting of 10 sentences in the practice test, and one list consisting of 20 sentences in the main test to estimate SNR required for 50% correct performance (i.e., correct repetition of 50% of the sentences). The HINT took about 10 min per participant to complete.

Cognitive Tests

Reading Span Test. The reading span test was designed to measure working memory capacity. The task requires the retention and recall of words while reading simple sentences. Baddeley et al. (1985) developed one such test based on the technique devised by Daneman and Carpenter (1980) in which sentences are presented visually, word by word, on a computer screen.

Several small lists of short sentences were presented to participants on the screen. Each sentence had to be judged as to its semantic correctness. Half of the sentences were semantically correct, and the other half were not (e.g., "Pappan kramade dottern"; "The father hugged his daughter" or "Räven skrev poesi"; "The fox wrote poetry") (Rönnberg et al., 1989; Rönnberg, 1990). The test began with two-sentence sets, followed by three-sentence sets, and so forth, up to five-sentence sets. Initially, participants were asked to press the "L" key if the sentence made sense or the "S" key for illogical sentences. After the set had been presented, participants were then required to recall either the first or final words of each sentence (e.g., "Pappan" and "Räven"; or "dottern"; and "poesi"), in the correct serial presentation order. Participants had about 3 s to press the "L" or "S" keys before the next sentence appeared. The computer instructed the participants to repeat either the first words or the last words of each sentence in the current set by typing them. The reading span score for each participant was equivalent to the total number of correctly recalled words across all sentences in the test, with a maximum score of 24. The reading span test took about 15 min per participant to complete.

The Paced Auditory Serial Addition Test (PASAT). The PASAT was initially designed to estimate information processing speed (Gronwall, 1977), but it is widely considered a test of attention (for a review, see Tombaugh, 2006). The task requires subjects to listen to a series of numbers (1–9), and to add consecutive pairs of numbers as they listen. As each number is presented, subjects must add that number to the previous number. For example, the following sequence of numbers is presented, one number at a time, every 2 or 3 s: 2, 5, 7, 4, and 6. The answers are: 7, 12, 11, and 10. The test demands a high level of attention, particularly if the numbers are presented quickly. In this study, we used a version of the PASAT in which digits were presented at an interval of

either 2 or 3 s (Rao et al., 1991), referred to as the PASAT 2 and the PASAT 3, respectively. Participants started with the PASAT 3, followed by the PASAT 2, with a short break between the two tests. The total number of correct responses (maximum possible = 60) at each pace was computed. The PASAT took about 15 min per participant to complete.

Preparation of gating tasks and procedure

A female speaker with clear enunciation and standard Swedish dialect read all of the items with normal intonation at a normal speaking rate in a quiet studio. Each item (consonant, word, or sentence) was recorded several times. We selected the item with the most natural intonation and clearest enunciation. Items were matched for sound level intensity. The sampling rate of the recording was 48 kHz, and the bit depth was 16 bits.

The onset and offset times of each recorded stimulus were marked in order to segment different types of stimuli. For each target, the onset time of the target was located as precisely as possible by inspecting the speech waveform (with Sound Studio 4 software) and using auditory feedback. The onset time was defined as the point where the signal amplitude ascended from the noise floor, according to the spectrograms in the Sound Studio 4 software. Each segmented section was then edited, verified, and saved as a ".wav" file. The gated stimuli were checked to eliminate click sounds. The root mean square value was computed for each stimulus waveform, and the stimuli were then rescaled to equate amplitude levels across the stimuli. A steady-state broadband noise, from Hällgren et al. (2006), was resampled and spectrally matched to the speech signals for use as background noise. The onset and offset of noise were simultaneous to the onset and offset of the speech signals.

The participants were tested individually in a quiet room. They were seated at a comfortable distance from a MacBook Pro (with Mac OS 10.6.7). Matlab (R2010b) was used to gate and present the stimuli binaurally through headphones (Sennheiser HDA200).

Participants received written instructions about the conditions for the different tasks (consonants, words, and final words in sentences), and performed several practice trials. In the practice trial, the sound level of the presentation was adjusted individually for each participant to a comfort level (approximately 60–65 dB). This sound level was used with no change in adjustment for that participant in both silent and noise conditions. In the noise condition (steady-state noise), the SNR was set at 0 dB, which was based on the findings of a pilot study using the current set of stimuli. During the practice session, the experimenter demonstrated how to use the keyboard to respond during the actual test. The participants were told that they would hear only part of a spoken target and would then hear progressively more. Participants were told to attempt identification after each presentation, regardless of how unsure they were about the identification of the stimulus, but to avoid random guessing. The participants were instructed to respond aloud and the experimenter recorded their responses. When necessary, the participants were asked to clarify their responses. The presentation of gates continued until the target was correctly identified on six consecutive presentations. If the target was not correctly identified, then the presentation continued until the entire target was disclosed, even if six or more consecutive responses were identical. Then, the experimenter started the next trial. When a target was not identified correctly, even after the whole target had been presented, its total duration plus one gate size was used as an estimate of the IP (cf. Elliott et al., 1987; Walley et al., 1995; Metsala, 1997; Hardison, 2005; Moradi et al., 2013). The rationale for this estimated IP was based on the fact that it was possible for participants to give correct responses at the last gate of a given target; hence, calculating an IP equal to the total duration of that target for two correct responses (even when late) and wrong responses would not be appropriate. No specific feedback was given to participants at any time during the session, except for general encouragement. Furthermore, there was no time pressure for responding to what was heard.

Each subject performed all of the gating tasks (consonants, words, and final words in sentences) in one session. All participants started with the identification of consonants task, followed by words task, and ended with the final words in sentences task. The type of condition (silence or noise) was counterbalanced across participants, such that half of the participants started with consonant identification in silence and then proceeded to consonant identification in noise, and vice versa for the other half of the participants. The order of items within each type of stimulus material (consonants, words, and sentences) varied between participants.

The full battery of gating tasks took 100–120 min per participant to complete. All of the tasks were performed in one session, but short rest periods were included to prevent fatigue. In the second session, the HINT, the reading span test, and the PASAT were administered. The order of the tests was counterbalanced across the participants. The second session took about 40 min per participant to complete.

RESULTS

GATING SPEECH TASKS

Figure 1 shows the mean IPs of consonants presented in both silence and noise conditions. Appendices A and B are confusion matrices for the 18 Swedish consonants presented in silence and noise, respectively. The values in the confusion matrices were extracted from correct and incorrect responses across all gates in the consonant gating paradigm tasks performed either in silence and noise. **Figure 2** shows the mean IPs for the gated speech tasks in both silence and noise conditions.

A Two-Way repeated-measure analysis of variance (ANOVA) was conducted to compare the mean IPs of the gated tasks (consonants, words, final words in LP sentences, and final words in HP sentences) in silence and noise. The results showed a main effect of the listening condition, $F_{(1,\ 20)} = 213.54$, $p < 0.001$, $\eta p^2 = 0.91$; a main effect of the gated tasks, $F_{(1.23,\ 24.54)} = 909.27$, $p < 0.001$, $\eta p^2 = 0.98$; and an interaction between listening condition and gated tasks, $F_{(1.58,\ 31.58)} = 49.84$, $p < 0.001$, $\eta p^2 = 0.71$. Four planned comparisons showed that the mean IPs of *consonants* in silence ($M = 101.78$, $SD = 11.47$) occurred earlier than in noise ($M = 166.14$, $SD = 26.57$), $t_{(20)} = 12.35$, $p < 0.001$, $d = 3.20$. In addition, the mean IPs of *words* in silence ($M = 461.97$, $SD = 28.08$) occurred earlier than in noise ($M = 670.51$, $SD = 37.64$), $t_{(20)} = 17.73$, $p < 0.001$, $d = 5.49$. The mean IPs of *final words in LP sentences* in silence ($M = 124.99$, $SD = 29.09$) were earlier than in noise ($M = 305.18$, $SD = 121.20$), $t_{(20)} = 7.67$, $p < 0.001$, $d = 2.56$. In addition, the

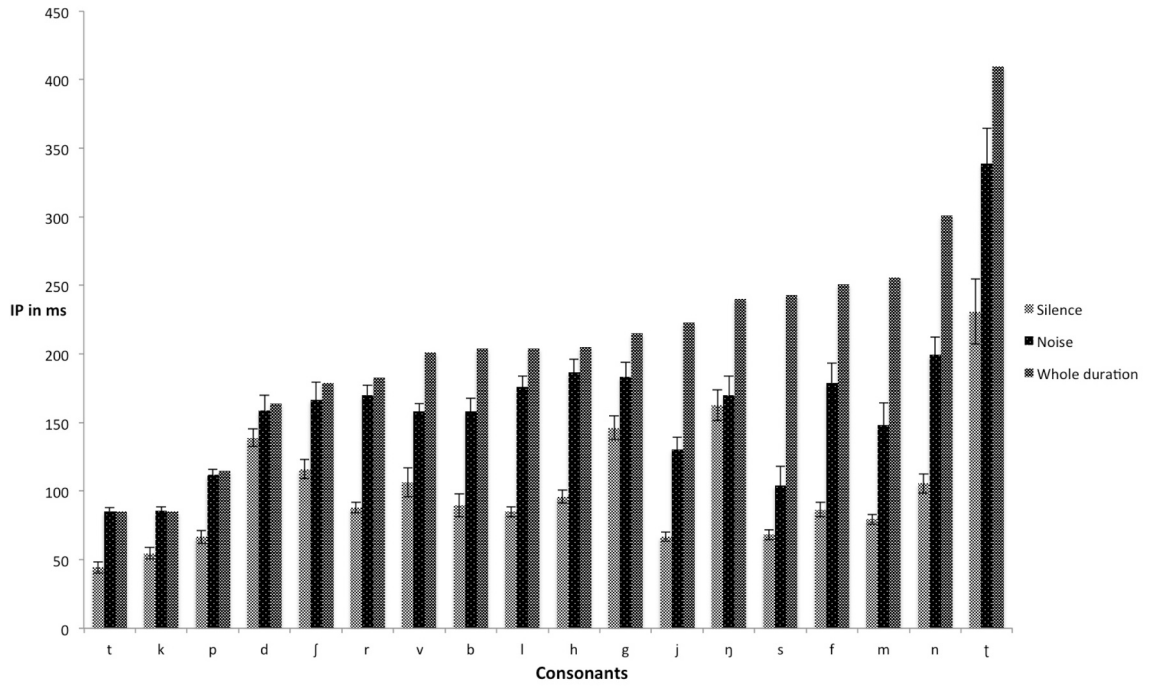


FIGURE 1 | Mean IPs (ms) for consonants in both silence and noise (with accompanying standard errors). IP, isolation point.

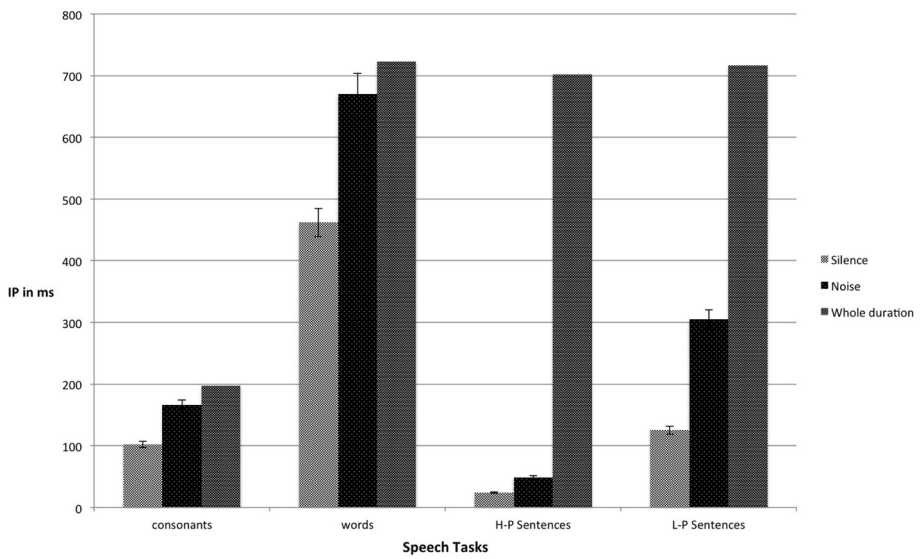


FIGURE 2 | Mean IPs (ms) for consonants, words, and final words in HP and LP sentences, in both silence and noise (with accompanying standard errors). IP, isolation point; HP, highly predictable; LP, low predictable.

mean IPs of *final words in HP sentences* in silence ($M = 23.96$, $SD = 3.31$) occurred earlier than in noise ($M = 48.57$, $SD = 23.01$), $t_{(20)} = 4.96$, $p < 0.001$, $d = 1.43$. We also analyzed our data by including only correct responses. The results showed that the mean IPs for consonants were 98.26 ($SD = 7.98$) ms in silence and 137.83 ($SD = 21.95$) ms in noise. In words, the mean IPs in silence were 456.31 ($SD = 21.49$) ms in silence and 505.89 ($SD = 50.77$) ms in noise. In final words in LP sentences, the mean IPs were 102.18 ($SD = 20.86$) ms in silence and 114.94 ($SD = 22.03$) ms in noise. In final words in HP sentences, the mean IPs were 23.86 ($SD = 3.33$) ms in silence and 42.24 ($SD = 15.24$) ms in noise. When comparing the results from two methods of IP calculations (i.e., including error responses with whole IPs of target stimuli plus one gate size, vs. including correct responses

only), there were subtle differences between IPs in silence; but greater differences in noise. For instance, when the IP calculation was based on correct responses only, the mean IPs for final word identification in sentences was 102.18 ms in silence and 114.94 ms in noise. However, when considering both correct and incorrect responses in the calculation of IPs for final word identification in sentences, the mean IPs became 124.99 ms in silence and 305.18 ms in noise. We therefore argue that the inclusion of error responses actually responses highlighted the interaction between noise and stimulus predictability (i.e., lexical, sentential, and semantic context), and that the interaction was logical and valid. In addition, the ANOVA on IPs only including correct responses showed the same pattern of results. There was a main effect of listening condition, $F_{(1,\ 20)} = 45.89$, $p < 0.001$, $\eta p^2 = 0.70$; a main effect of the gated tasks, $F_{(1.68,\ 33.49)} = 3545.27$, $p < 0.001$, $\eta p^2 = 0.99$; and an interaction between listening condition and gated tasks, $F_{(1.55,\ 30.91)} = 6.10$, $p < 0.01$, $\eta p^2 = 0.23$.

Table 1 reports the percentage of correct responses for each of the gated tasks performed in both silence and noise conditions. A Two-Way repeated-measures analysis (ANOVA) showed a main effect of listening condition, $F_{(1,\ 20)} = 223.41$, $p < 0.001$, $\eta p^2 = 0.92$; a main effect of the gated tasks, $F_{(3,\ 60)} = 36.86$, $p < 0.001$, $\eta p^2 = 0.65$; and an interaction between listening condition and gated tasks, $F_{(3,\ 60)} = 33.24$, $p < 0.001$, $\eta p^2 = 0.62$. Four planned comparisons showed that noise reduced the accuracy for the identification of consonants, $t_{(20)} = 7.50$, $p < 0.001$, $d = 2.21$; words, $t_{(20)} = 15.14$, $p < 0.001$, $d = 4.26$; final words in LP sentences, $t_{(20)} = 4.28$, $p < 0.001$, $d = 1.10$; and final words in HP sentences, $t_{(20)} = 2.90$, $p < 0.009$, $d = 1.51$.

CORRELATIONS BETWEEN GATING SPEECH TASKS, THE HINT, AND THE COGNITIVE TESTS

Table 2 shows the means responses of participants for the HINT, PASAT 3, PASAT 2, and the reading span test. The correlation matrix (**Table 3**) shows the Pearson correlations between the IPs of gated tasks in both silence and noise conditions (lower scores in the gated tasks reflect better function), the HINT scores (lower scores in the HINT reflect better function), and the reading span test and PASAT scores (higher scores in the reading span test and PASAT reflect better function). The PASAT 2 scores were significantly correlated with the HINT scores, the reading span test scores, IPs of consonants in noise, and IPs of words in noise. This finding suggested that lower IP scores for consonants and words in noise were usually associated with better performance on the HINT and PASAT 2. The reading span test scores were also significantly correlated with the HINT scores and IPs for consonants in noise, indicating that better performance on the reading span test was associated with better performance on the HINT and earlier IPs for consonants in noise. The HINT scores were significantly correlated with IPs for consonant and word identification in noise; the better the listeners performed on the HINT, the earlier they generally identified consonants and words in noise.

We also compared pairs of correlational coefficients in silence and noise (**Table 4**). The results showed that three pairwise correlations were significantly different from each other. We also tested if there is a difference between the means of the correlation coefficients of the two matrices (between the IPs and the scores of the cognitive tasks and the HINT, with z transformed correlation coefficients). We therefore first put all correlation coefficients in the same (logical) direction. Then we tested the means difference with a paired two-tailed t test. In this case, $n = 12$, since we used the number of paired correlations as "individuals." The result was $t_{(10)} = 3.64$, $p = 0.005$, $d = 1.05$, that is, a significant difference between the mean correlation coefficients for silence versus noise, with a large effect size. We argue that the data pattern, comparing correlations for the silent versus noisy conditions, shows a valid difference such that cognitive tests are generally more strongly correlated with IPs for consonants and words in the noisy conditions compared to the silent conditions. Thus, support for the validity of this conclusion comes from (a) the overall qualitative pattern of differences in correlation matrices, (b) from inferential statistics comparing pairwise correlations, and (c) from statistical comparison of the entire (pooled) correlation matrices.

DISCUSSION

HOW DOES NOISE GENERALLY AFFECT IPS?

The results show that noise generally delayed the IPs for the identification of consonants, words, and final words in LP and HP sentences, which is in line with the predictions. Furthermore, our results demonstrate the advantage of IPs over accuracy especially in the silent condition. While there was a ceiling effect for identification of consonants, words, and final words in HP sentences in silence (over 95% correct responses), there was substantial variation in their IPs.

HOW DOES NOISE AFFECT IPS WHEN CONSIDERING LINGUISTIC (i.e., LEXICAL AND SENTENTIAL) CONTEXT?

Consonants

There was variation in the IPs of consonants, implying that the location of critical cues for their identification varies across

Table 1 | Identification accuracy for gating spoken stimuli.

Type of gated stimuli	Silence mean (*SD*)	Noise mean (*SD*)
Consonants	97.4 (3.8)	70.1 (17.5)
Words	96.3 (5.2)	34.6 (17.1)
HP Sentences	94.8 (7.7)	85.7 (8.0)
LP Sentences	87.3 (7.3)	67.1 (20.3)

SD, standard deviation; HP, highly predictable; LP, low predictable.

Table 2 | HINT, PASAT 3, PASAT 2, and reading span test results.

Type of task	Mean (SD)
HINT	−3.1 (1.2)
PASAT 3	51.2 (4.4)
PASAT 2	40.0 (6.2)
Reading span test	21.6 (1.7)

HINT, Hearing in Noise Test; PASAT, Paced Auditory Serial Attention Test (digits are presented at an interval of 2 or 3 s); SD, standard deviation.

Table 3 | Correlation matrix for gating speech variables, HINT, and cognitive test results.

	1	2	3	4	5	6	7	8	9	10	11	12
1. HINT		−0.09	−0.63**	−0.58**	0.27	0.73**	−0.26	0.58**	0.08	0.24	0.00	0.22
2. PASAT 3			0.51*	0.55*	−.012	−0.22	0.06	0.07	0.04	−0.14	−0.23	−0.39
3. PASAT 2				0.65**	−0.39	−0.68**	0.22	−0.51*	0.00	−0.21	0.03	−0.34
4. RST					−0.19	−0.51*	0.23	−0.30	−0.21	−0.41	−0.35	−0.42
5. Consonant-S						0.44*	−0.09	0.36	−0.15	0.03	0.07	0.32
6. Consonant-N							−0.03	0.56**	0.18	0.35	0.24	0.34
7. Word-S								−0.33	0.20	−0.11	−0.11	−0.27
8. Word-N									0.16	0.27	−0.16	0.16
9. HP-S										0.33	0.15	−0.04
10. LP-S											0.50*	0.56**
11. HP-N												0.58**
12. LP-N												

HINT, Hearing in Noise Test; PASAT, Paced Auditory Serial Attention Test (digits are presented at an interval of 2 or 3 s); RST, Reading Span Test; Consonant-S, gated consonant identification in silence; Consonant-N, gated consonant identification in noise; Word-S, gated word identification in silence; Word-N, gated word identification in noise; HP-S, gated final word identification in highly predictable sentences in silence; LP-S, gated final word identification in low predictable sentences in silence; HP-N, gated final word identification in high predictable sentences in noise; LP-N, gated final word identification in low predictable sentences in noise. $^{*}p < 0.05$. $^{**}p < 0.01$.

Table 4 | Fisher's Z scores to compare correlation coefficients between silence and noise.

	Consonants	Words	Final words in HP	Final words in LP
HINT	−2.69*	−2.69*	0.26	0.09
PASAT 3	0.91	−0.03	1.02	1.23
PASAT 2	1.56	2.18*	−0.11	0.62
Reading span test	1.48	1.46	0.55	0.08

$^{*}p < 0.05$.

consonants, corroborating the findings of Smits (2000). For instance, the time ratio in silence showed that /b f h j l m n s/ required roughly one-third and /d k p ʃ/ required about two-thirds of their full durations for identification. Noise extended the amount of time required for correct identification of consonants. Consonants in the noise condition required longer exposure to be identified because their critical features were masked. In our study, the accuracy rate for correct identification of consonants was about 97% in silence, which dropped to 70% in noise (**Table 1**). This is consistent with the findings of Apoux and Healy (2011), wherein listeners correctly identified 68% of consonants in speech-shaped noise at 0 dB SNR. Cutler et al. (2008) reported about 98% correct identification of consonants in quiet conditions, and about 80% in eight-talker babble noise. In addition, the results in the confusion matrix (Supplementary meterials) for identification of Swedish consonants show that at 0 SNR dB, /b d g h k r ʈ ʃ t/ are often confused with each other, /f l m ŋ p r/ are moderately confused with each other, and /j n s/ hardly ever confused with each other.

Words

Noise also increased the amount of time required for the correct identification of Swedish monosyllabic words. In silence, just over half of the duration of a word was required for identification. This finding is consistent with previous studies using English words. Grosjean (1980) showed that about half of the segments of words were required for word identification. In noise, almost the full duration of words was required for identification in the current study. **Table 3** shows that consonant identification in noise was significantly correlated with word identification in noise and HINT performance, which might imply that the misperception of a consonant was misleading for the identification of words in noise. In fact, the incorrect identification of just one consonant or vowel (in consonant-vowel-consonant word format) can lead to the activation of another candidate in the lexicon, and realizing the misperception and finding another candidate takes more time. In summary, noise delays word identification and increases the risk of misidentification, and may make it impossible to identify a word at all. This was also the case in the present study. Not only were the IPs delayed by noise, accuracy was also impeded: about 96% accuracy in silence versus 35% in noise (see **Table 1**). These results are also consistent with previous studies (Chermak and Dengerink, 1981; Studebaker et al., 1999).

Final words in sentences

The presence of noise delayed final word identification in LP and HP sentences. In silence, highly relevant contextual information seems to prohibit the activation of other lexical candidates even earlier than word-alone presentation. However, the presence of noise resulted in delayed identification of final words even in both LP and HP sentences. These results are in agreement with Aydelott and Bates (2004) who reported that the perceptual clarity of speech signal impacts on the ability to make use of semantic context to aid in lexical processing. They studied how response times to target words in congruent sentences were influenced by low-pass filtering of prior context. Their result showed that low-pass filtering reduced the facilitation of semantic context on identification of target words. The mean IPs for

final-word identification in LP sentences (125 ms in silence and 305 ms in noise) were found to be even shorter than the mean IPs for isolated words in silence (462 ms), demonstrating that even low predictable information can speed up decoding of the speech signal (cf. Salasoo and Pisoni, 1985; Van Petten et al., 1999). The accuracy rates for final words in HP and LP sentences in noise were 86 and 67%, respectively, which also is consistent with Kalikow et al. (1977). As **Table 1** shows, accuracy in the noise condition was higher for final words in LP sentences (67%) than for the identification of isolated words (35%). We assume that (similar to the identification of isolated words) masking consonants with noise activates other consonants which form words that are still related to the contents of LP sentences, and eliminating them is time consuming. However, because there is *some* contextual information in LP sentences that excludes *some* candidates in the mental lexicon, correct identification of final words in LP sentences is accomplished at earlier gates compared to the identification of words in isolation (cf. Ladefoged and Broadbent, 1957).

To conclude, the results from comparing IPs from gated speech stimuli in silence versus noise suggest that less information is available in noise because of masking (e.g., Dorman et al., 1998; Shannon et al., 2004; for a review, see Assmann and Summerfield, 2004). We suppose that the combination of noise with speech stimuli hindered the listener from accessing the detailed acoustic information (in particular for consonants and words), whereas this access to the detailed acoustic information was readily available in a silent condition. As a consequence, noise delays the amount of time required (in other words, necessitates more acoustic information) for correct identification of speech stimuli to occur. In addition, our finding is in agreement with the "active sensing" hypothesis (for a review see Zion Golumbic et al., 2012) which suggests that the brain consistently makes predictions about the identity of the forthcoming stimuli, rather than passively waiting to receive and thereafter identify the stimuli (Rönnberg et al., 2013).

COGNITIVE DEMANDS OF SPEECH PERCEPTION IN SILENCE AND NOISE

HINT

Results showed that HINT performance was correlated with measures of working memory capacity (the reading span test), and attention capacity (PASAT 2). Listeners with better hearing-in-noise ability had higher scores in the tests of working memory and attention capacities. This result corroborates the previous studies that reported correlations between sentence comprehension in noise and the reading span test (e.g., Rudner et al., 2009; Ellis and Munro, 2013). Successful performance in the HINT requires filtering out the noise as well as focusing on the target signal, temporarily storing all of the words within sentences, and remembering them. It is therefore reasonable that HINT performance is correlated with the measures of attention and working memory capacities. One of the reasons for this correlation can be found in neuroimaging studies that demonstrate that the activation of auditory (superior temporal sulcus and superior temporal gyrus) *and* cognitive (e.g., left inferior frontal gyrus) brain areas are provoked during the comprehension of degraded sentences compared to clear speech (Davis et al., 2011; Wild et al., 2012; Zekveld et al., 2012). According to Giraud and Price (2001) and Indefrey and Cutler (2004), the tasks that require extra cognitive processes, such as attention and working memory, activate prefrontal brain areas that include the inferior frontal gyrus. Both stimulus degradation (Wild et al., 2012) and speech-in-noise seem to call on similar neurocognitive substrates (Zekveld et al., 2012). Thus, the observed HINT correlations are in agreement with previous studies.

Consonants

Better performance in the HINT, reading span test, and PASAT were associated with earlier identification of consonants in noise. Neuroimaging studies have also revealed that ambiguous phoneme identification requires top-down cognitive support from prefrontal brain areas in addition to predominantly auditory brain areas to correctly identify ambiguous phonemes (Dehaene-Lambertz et al., 2005; Dufor et al., 2007). However, our finding is not in agreement with Cervera et al. (2009) who showed no significant correlations between tests of working memory capacity (serial recall and digit ordering) and consonant identification in noise at 6 dB SNR. One explanation for this inconsistency may be the fact that we presented the gated consonants at 0 dB SNR, which is more difficult and cognitively demanding than the task used by Cervera et al. (2009).

Words in isolation

There was a significant correlation between the IPs of words in noise and scores for the HINT and PASAT 2, suggesting that listeners with better attention capacity and hearing-in-noise abilities identified words in noise earlier than those with poorer abilities. Shahin et al. (2009) degraded words by inserting white noise bursts around the affricatives and fricatives (of words). They found greater activation of the left inferior frontal gyrus during the processing of degraded words, which they suggested was implicated to "repair" the illusion of hearing words naturally when in reality participants had heard degraded words. In our study, it can be concluded that listeners who had better hearing-in-noise and attention capacities were able to repair this "illusion of hearing words naturally" earlier than those with poorer abilities, which resulted in shorter IPs for words in noise. It should be noted that we expected to see a significant correlation between IPs for words in noise and also with the reading span test (working memory capacity). However, there was no significant relationship between IPs for words in noise and test of working memory capacity. One explanation might be that for word identification, we presented the first phoneme of the words and then started the gating paradigm from the second phoneme (in a consonant-vowel-consonant format). In addition, the gate size for word identification was twice as large as for consonants. We therefore assume that this procedure for word identification reduced the demand on working memory for identification of words in noise. With the advantage of hindsight, this potentially important procedural detail should be accounted for in future gating research.

Overall, our findings for the identification of consonants and words in silence and noise are consistent with general predictions

of the ELU model (Rönnberg et al., 2008, 2013), which suggests that speech perception is mostly effortless under optimum listening conditions, but becomes effortful (cognitively demanding) in degraded listening conditions. Clearly audible signals may not depend as much on working memory and attentional capacities, because they can be implicitly and automatically mapped onto the phonological representations in the mental lexicon.

Final words in sentences

Our results showed that there were no correlations between the IPs for final words in HP and LP sentences in noise condition and measures of working memory and attention. This finding is consistent with some previous studies which have shown that when listening is challenged by noise, prior contextual knowledge acts as a major source of disambiguation by providing expectations about which word (or words) may appear at the end of a given sentence (cf. Cordillo et al., 2004; Obleser et al., 2007). Hence, it can be assumed that at an equal SNR, the identification of final words in sentences is easier than the identification of consonants and words uttered in isolation; the sentence context makes final word identification less cognitively demanding (i.e., less effortful) than the identification of isolated consonants and words. This result is not in agreement with the original version of the ELU model (Rönnberg, 2003; Rönnberg et al., 2008) in which there was no postulated mechanism for the contextual elimination of lexical candidates. However, in the recent updated version of the ELU model (Rönnberg et al., 2013), the early top-down influence of semantic context on speech recognition under adverse conditions is taken into account. The model suggests that because of the combined semantic and syntactic constraints in a given dialog, listeners may need little information regarding a target signal, if the preceding contextual priming is sufficiently predictive.

In our study, while there were correlations between measures of cognitive tests and the HINT, no significant correlations were observed between cognitive tests and the IPs of final words in (LP and HP) sentences. One possible explanation might be that performance on the HINT requires listeners to remember *all* of the words in each sentence correctly, at varying SNRs, which taxes working memory (Rudner et al., 2009; Ellis and Munro, 2013). Successful performance in this task requires the short-term decoding and maintenance of masked speech stimuli, and the subsequent retrieval of the whole sentence. However, the identification of final words in sentences simply requires the tracking of incoming speech stimuli, and the subsequent guessing of the final words is based on the sentential context and the first consonant of the final word. This prior context plus initial consonant is likely to reduce cognitive demands, which was presumably lower than that required for the HINT performance. In addition, performance in the HINT was based on 50% correct comprehension of sentences in noise. As **Table 1** shows, the mean accuracy rates in the noise condition for final words in LP and HP sentences were about 67 and 86%, respectively, which are higher than the 50% correct comprehension rate for sentences in the HINT. Furthermore, the mean SNR for HINT performance in the present study was −3.1 dB (**Table 2**), while final words in sentences in noise condition were presented at 0 dB. Thus, it can be concluded that identification in the LP and HP sentences under the noise condition was easier than HINT identification, and as such tapped into the implicit mode of processing postulated by the ELU model. Future studies are needed in order to investigate the correlations between tests of working memory and attention and IPs for final-word identification in sentences at lower SNRs. It is likely that by decreasing the SNR, the demand on working memory and attention capacities will increase even for such sentence completion tasks.

In our study, the PASAT demonstrated a significant correlation with the reading span test, which is in agreement with previous studies (Sherman et al., 1997; Shucard et al., 2004). Interestingly, only the PASAT 2 was correlated with HINT performance and consonant and word identification in noise, whereas the PASAT 3 was not. This probably suggests that the significant relationship with speech perception in noise was related to the attention-demanding aspect of the task, because PASAT 2 is more paced and taxing. This result is in line with the review by Akeroyd (2008), who argued that only sufficiently taxing cognitive tasks are correlated with speech perception in degraded listening conditions. In Akeroyd (2008), not all cognitive tests yielded significant correlations with noise; only specific measures of cognitive abilities such as working memory (e.g., the reading span test) were correlated with speech-in-noise tasks, whereas general, composite, cognitive measures (like IQ) were not.

Taken together, noise delays the IPs for identification of speech stimuli. In addition, the results suggest that early and correct identification of spoken signals in noise requires an interaction between auditory, cognitive, and linguistic factors. Speech tasks that lack a contextual cue, such as consonants and words presented in isolation, more probably draw on the interaction between auditory and explicit cognitive factors. However, when the perception of speech in noise relies on prior contextual information, or when there is no noise, superior auditory and cognitive abilities are less critical.

CONCLUSIONS

The identification of consonants, words, and final words in sentences was delayed by noise. The mean correlation between cognitive tests and IPs was stronger for the noisy condition than for the silent condition. Better performance in the HINT was correlated with greater capacities of working memory and attention. Rapid identification of consonants in noise was associated with greater capacities of working memory and attention and also HINT performance; and rapid identification of words in noise was associated with greater capacity of attention and HINT performance. However, the identification of final words in sentences in the noise condition was not demanding enough to depend on working memory and attentional capacities to aid identification. This is presumably due to the facilitation from prior sentential context, lowering the demands on explicit cognitive resources.

ACKNOWLEDGMENTS

Part of data from the present study has been used in Moradi et al. (2013) in order to compare audiovisual versus auditory gating presentation on IPs and accuracy of speech perception. This research was supported by the Swedish Research Council

(349-2007-8654). The authors would like to thank Katarina Marjanovic for speaking the recorded stimuli and two reviewers for their insightful comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00531/abstract

REFERENCES

Akeroyd, M. A. (2008). Are individual differences in speech perception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. Int. *J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Apoux, F., and Healy, E. W. (2011). Relative contribution of target and masker temporal fine structure to the unmasking of consonants in noise. *J. Acoust. Soc. Am.* 130, 4044–4052. doi: 10.1121/1.3652888

Assmann, P. F., and Summerfield, A. Q. (2004). "The perception of speech under adverse conditions," in *Speech Processing in the Auditory System*, Vol. 14. *Springer Handbook of Auditory Research*, eds S. Greenberg, W. A. Ainsworth, A. N. Popper and R. Fay (New York, NY: Springer-Verlag), 231–308.

Aydelott, J., and Bates, E. (2004). Effects of acoustic distortion and semantic context on lexical access. *Lang. Cogn. Proc.* 19, 29–56. doi: 10.1080/01690960344000099

Baddeley, A. D., Logie, R., Nimmo-Smith, I., and Brereton, R. (1985). Components of fluent reading. *J. Mem. Lang.* 24, 119–131. doi: 10.1016/0749-596X(85)90019-1

Benichov, J., Cox, L. C., Tun, P. A., and Wingfield, A. (2012). Word recognition within a linguistic context: effects of age, hearing acuity, verbal ability, and cognitive function. *Ear. Hear.* 32, 250–256. doi: 10.1097/AUD.0b013e31822f680f

Boehnke, S. E., and Phillips, D. P. (1999). Azimuthal tuning of human perceptual channels for sound location. *J. Acoust. Soc. Am.* 106, 1948–1955. doi: 10.1121/1.428037

Carlyon, R. P., Cusack, R., Foxton, J. M., and Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 115–127. doi: 10.1037/0096-1523.27.1.115

Cervera, T. C., Soler, M. J., Dasi, C., and Ruiz, J. C. (2009). Speech recognition and working memory capacity in young-elderly listeners: effects of hearing sensitivity. *Can. J. Exp. Psychol.* 63, 216–226. doi: 10.1037/a0014321

Chermak, G. D., and Dengerink, J. (1981). Word identification in quiet and in noise. *Scand. Audiol.* 10, 55–60. doi: 10.3109/01050398109076162

Colflesh, G. J. H., and Conway, A. R. A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychon. Bull. Rev.* 14, 699–703. doi: 10.3758/BF03196824

Conway, A. R. A., Cowan, N., and Bunting, M. F. (2001). The cocktail party phenomenon revisited: the importance of WM capacity. *Psychon. Bull. Rev.* 8, 331–335. doi: 10.3758/BF03196169

Cordillo, E. R., Aydelott, J., Matthews, P. M., and Devlin, J. T. (2004). Left inferior prefrontal cortex activity reflects inhibitory rather than faciliatory priming. *J. Cogn. Neurosci.* 16, 1552–1561. doi: 10.1162/0898929042568523

Cutler, A., Garcia Lecumberri, M. L., and Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: effects of local context. *J. Acoust. Soc. Am.* 124, 1264–1268. doi: 10.1121/1.2946707

Cutler, A., Smits, R., and Cooper, N. (2005). Vowel perception: effects of non-native language vs. non-native dialect. *Speech Commun.* 47, 32–42. doi: 10.1016/j.specom.2005.02.001

Dai, H., Scharf, B., and Buss, S. (1991). Effective attenuation of signal in noise under focused attention. *J. Acoust. Soc. Am.* 89, 2837–2842. doi: 10.1121/1.400721

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal. Learning. Verbal. Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Davis, M. H., Ford, M. A., Kherif, F., and Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved spare fMRI. *J. Cogn. Neurosci.* 23, 3914–3932. doi: 10.1162/jocn_a_00084

Davis, M. H., and Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132–147. doi: 10.1016/j.heares.2007.01.014

Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger–Charolles, L., Jobet, A., and Dehane, S. (2005). Neural correlates of switching from auditory to speech perception. *Neuroimage* 24, 21–33. doi: 10.1016/j.neuroimage.2004.09.039

De Neys, W., Schaeken, W., and D'Ydewalle, G. (2003). "Working memory span and everyday conditional reasoning: a trend analysis," in *Proceeding of the twenty-fifth Annual Conference of the Cognitive Science Society* (Boston, MA), 312–317.

Dorman, M. F., Loizou, P. C., Fitzke, J., and Tu, Z. (1998). The recognition of sentences in noise by normal hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels. *J. Acoust. Soc. Am.* 104, 3583–3585. doi: 10.1121/1.423940

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2005). Word recognition in noise at higher-than-normal levels: decreases in scores and increases in masking. *J. Acoust. Soc. Am.* 118, 914–922. doi: 10.1121/1.1953107

Dufor, O., Serniclaes, W., Sprenger–Charolles, L., and Démonet, J.-F. (2007). Top-down processes during auditory phoneme categorization in dyslexia: a PET study. *Neuroimag* 34, 1692–1707. doi: 10.1016/j.neuroimage.2006.10.034

Elliott, L. L., Hammer, M. A., and Evan, K. E. (1987). Perception of gated, highly familiar spoken monosyllabic nouns by children, teenagers, and older adults. *Percept. Psychophys.* 42, 150–157. doi: 10.3758/BF03210503

Ellis, R. J., and Munro, K. J. (2013). Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *Int. J. Audiol.* 52, 14–22. doi: 10.3109/14992027.2012.721013

Etikregler för humanistisk-samhällsvetenskaplig forskning. (n.d.). *Forskningsetiska principer i humanistisk-samhällsvetenskaplig forskning. Antagna av Humanistisk-samhällsvetenskapliga forskningsrådet i mars 1990, reviderad version April 1999.* [*Principles of Research Ethics for Research within the Arts and Social Sciences. Accepted by the Research Council for Arts and Sciences in March 1990, revised version April 1999.*] Available online at: http://www.stingerfonden.org/documents/hsetikregler.pdf (Accessed February 10, 2012).

Foo, C., Rudner, M., Rönnberg, J., and Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *J. Am. Acad. Audiol.* 18, 618–631. doi: 10.3766/jaaa.18.7.8

Giraud, A. L., and Price, C. J. (2001). The constraints functional neuroimaging places on classical models of auditory word processing. *J. Cogn. Neurosci.* 13, 754–765. doi: 10.1162/08989290152541421

Grant, K. W., and Seitz, P. F. (2000). The recognition of isolated words and words in sentences: individual variability in the use of sentence context. *J. Acoust. Soc. Am.* 107, 1000–1011. doi: 10.1121/1.428280

Gronwall, D. (1977). Paced Auditory serial addition test: a measure of recovery from concussion. *Percept. Mot. Skills* 44, 367–373. doi: 10.2466/pms.1977.44.2.367

Grosjean, F. (1980). Spoken word recognition processes and gating paradigm. *Percept. Psychophys.* 28, 267–283. doi: 10.3758/BF03204386

Hardison, D. M. (2005). Second-language spoken word identification: effects of perceptual training, visual cues, and phonetic environment. *Appl. Psychol.* 26, 579–596. doi: 10.1017/S0142716405050319

Hällgren, M., Larsby, B., and Arlinger, S. (2006). A Swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition. *Int. J. Audiol.* 45, 227–237. doi: 10.1080/14992020500429583

Indefrey, P., and Cutler, A. (2004). "Prelexical and lexical processing in listening" in *The Cognitive Neurosciences, 3rd Edn*, ed M. S. Gazzaniga (Cambridge, MA: MIT Press), 759–774.

Kalikow, D. N., Stevens, K. N., and Elliott, L. L. (1977). Development of speech intelligibility in noise using sentence materials with controlled word predictability. *J. Acoust. Soc. Am.* 61, 1337–1351. doi: 10.1121/1.381436

Kane, M. J., and Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 336–358. doi: 10.1037/0278-7393.26.2.336

Ladefoged, P., and Broadbent, D. E. (1957). Information conveyed by vowels. *J. Acoust. Soc. Am.* 29, 98–104. doi: 10.1121/1.1908694

Lidestam, B. (2014). Audiovisual presentation of video-recorded stimuli at a high frame rate. *Behav. Res. Methods.* 46, 499–516. doi: 10.3758/s13428-013-0394-2

Lively, S. E., Pisoni, D. B., and Goldinger, S. D. (1994). Spoken word recognition: research and theory, in *Handbook of Psycholinguistics*, ed M. A. Gernsbacher (San Diego, CA: Academic Press), 265–301.

Lunner, T., Rudner, M., and Rönnberg, J. (2009). Cognition and hearing aids. *Scand. J. Psychol.* 50, 395–403. doi: 10.1111/j.1467-9450.2009.00742.x

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition* 25, 71–102. doi: 10.1016/0010-0277(87)90005-9

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse listening conditions: a review. *Lang. Cogn. Proc.* 27, 953–978. doi: 10.1080/01690965.2012.705006

Mesgarani, N., and Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485, 233–236. doi: 10.1038/nature11020

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Mem. Cogn.* 25, 47–56. doi: 10.3758/BF03197284

Miller, G. A., Heise, G. A., and Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *J. Exp. Psychol.* 41, 329–335. doi: 10.1037/h0062491

Miller, G. A., and Isard, S. (1963). Some perceptual consequences of linguistic rules. *J. Verbal. Learning. Verbal. Behav.* 2, 217–228. doi: 10.1016/S0022-5371(63)80087-0

Miller, G. A., and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Am.* 27, 338–352. doi: 10.1121/1.1907526

Mondor, T. A., Zatorre, R. J., and Terrio, N. A. (1998). Constrains on the selection of auditory information. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 66–79. doi: 10.1037/0096-1523.24.1.66

Moradi, S., Lidestam, B., and Rönnberg, J. (2013). Gated audiovisual speech identification in silence vs. noise: effects on time and accuracy. *Front. Psychol.* 4:359. doi: 10.3389/fpsyg.2013.00359

Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). Development of the Hearing In Noise Test (HINT) for the measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. Am.* 95, 338–352. doi: 10.1121/1.408469

Oblerser, J., Wise, R. J. S., Dresner, X., and Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *J. Neurosci.* 27, 2283–2289. doi: 10.1523/JNEUROSCI.4663-06.2007

Parikh, G., and Loizou, P. C. (2005). The influence of noise on vowel and consonant cues. *J. Acoust. Soc. Am.* 118, 3874–3888. doi: 10.1121/1.2118407

Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., and Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nat. Neurosci.* 7, 658–663. doi: 10.1038/nn1256

Phatak, S. A., and Allen, J. B. (2007). Consonant and vowel confusions in speech-weighted noise. *J. Acoust. Soc. Am.* 121, 2312–2316. doi: 10.1121/1.2642397

Poeppel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071–1086. doi: 10.1098/rstb.2007.2160

Rabbitt, P. M. (1968). Channel-capacity, intelligibility, and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158

Rao, S. M., Leo, G. J., Bernardin, L., and Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis: frequency, patterns, and prediction. *Neurology*. 41, 685–691. doi: 10.1212/WNL.41.5.685

Röer, J. P., Bell, R., Dentale, S., and Buchner, A. (2011). The role of habituation and attentional orienting in the disruption of short-term memory performance. *Mem. Cogn.* 39, 839–850. doi: 10.3758/s13421-010-0070-z

Rönnberg, J. (1990). Cognitive and communicative functions: the effects of chronological age and "handicap age." *Euro. J. Cogn. Psychol.* 2, 253–273. doi: 10.1080/09541449008406207

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42, S68–S76. doi: 10.3109/14992020309074626

Rönnberg, J., Arlinger, S., Lyxell, B., and Kinnefors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive functions. *J. speech. Lang. Hear. Res.* 32, 725–735.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Sys. Neurosci.* 7:00031. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi: 10.1080/14992020802301167

Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12, 263–269. doi: 10.4103/1463-1741.70505

Rudner, M., Foo, C., Rönnberg, J., and Lunner, T. (2009). Cognition and aided speech recognition in noise: specific role for cognitive factors following nine-weeks experience with adjusted compression settings in hearing aids. *Scand. J. Psychol.* 50, 405–418. doi: 10.1111/j.1467-9450.2009.00745.x

Rudner, M., Rönnberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairments. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4

Salasoo, A., and Pisoni, D. (1985). Interaction of knowledge source in spoken word identification. *J. Mem. Lang.* 24, 210–231. doi: 10.1016/0749-596X(85)90025-7

Sawusch, J. R. (1977). Processing of place information in stop consonants. *Percept. Psychophys.* 22, 417–426. doi: 10.3758/BF03199507

Shahin, A. J., Bishop, C. W., and Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage* 44, 1133–1143. doi: 10.1016/j.neuroimage.2008.09.045

Shannon, R. V., Fu, Q.–J., and Galvin, J. (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta. Otolarygol. Suppl.* 552, 1–5. doi: 10.1080/03655230410017562

Sheldon, S., Pichora-Fuller, M. K., and Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *J. Acoust. Soc. Am.* 123, 489–499. doi: 10.1121/1.2783762

Sherman, E. M. S., Strauss, E., and Spellacy, F. (1997). Validity of the Paced Auditory Serial Addition Test (PASAT) in adults referred for neuropsychological assessment after head injury. *Clin. Neuropsychol.* 11, 34–45. doi: 10.1080/13854049708407027

Shinn-Cunningham, B. G., and Best, V. (2008). Selective attention in normal and impaired hearing. *Trends. Amplif.* 12, 283–299. doi: 10.1177/1084713808325306

Shucard, J. L., Parrish, J., Shucard, D. W., McCabe, D. C., Benedict, R. H. B., and Ambrus, J. (2004). Working memory and processing speed deficits in systemic lupus erythematous as measured by the Paced Auditory Serial Addition Test. *J. Int. Neuropsychol. Soc.* 10, 35–45. doi: 10.1017/S1355617704101057

Smits, R. (2000). Temporal distribution of information for human consonant recognition in VCV utterances. *J. Phon.* 27, 111–135. doi: 10.1006/jpho.2000.0107

Snyder, J. S., Alain, C., and Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *J. Cogn. Neurosci.* 18, 1–13. doi: 10.1162/089892906775250021

Sörqvist, P., and Rönnberg, J. (2012). Episodic long-term memory of spoken discourse masked by speech: what role for working memory capacity. *J. Speech. Lang. Hear. Res.* 55, 210–218. doi: 10.1044/1092-4388(2011/10-0353)

Sörqvist, P., Stenfelt, S., and Rönnberg, J. (2012). Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: toward a unified view of attention. *J. Cogn. Neurosci.* 24, 2147–2154. doi: 10.1162/jocn_a_00275

Språkbanken (the Swedish Language Bank). (2011). Available online at: http://spraakbanken.gu.se/ (Accessed November 30, 2011).

Studebaker, G. A., Sherbeco, R. L., McDaniel, D. M., and Gwaltney, C. A. (1999). Monosyllabic word recognition at higher-than-normal speech and noise levels. *J. Acoust. Soc. Am.* 105, 2431–2444. doi: 10.1121/1.426848

Taylor, W. L. (1953). "Cloze procedure": a new tool for measuring readability. *Jour. Q.* 30, 415–433.

Tombaugh, T. N. (2006). A comprehensive review of the Paced Auditory Serial Addition Test (PASAT). *Arch. Clin. Neuropsychol.* 21, 53–76. doi: 10.1016/j.acn.2005.07.006

Troille, E., Cathiard, M.-A., and Abry, C. (2007). "Consequences on bimodal perception of the timing of the consonant and vowel audiovisual flows," in *Proceedings of International Conference on Audio-Visual Speech Processing* (Hilvarenbeek), 281–286.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., and Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 394–417. doi: 10.1037/0278-7393.25.2.394

Walley, A. C., Michela, V. L., and Wood, D. R. (1995). The gating paradigm: effects of presentation format on spoken word recognition by children and adults. *Atten. Percept. Psychophys.* 57, 343–351. doi: 10.3758/BF03213059

Wang, M., and Bilger, R. (1973). Consonant confusions in noise: a study of perceptual features. *J. Acoust. Soc. Am.* 54, 1248–1266. doi: 10.1121/1.1914417

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., and Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *J. Neurosci.* 32, 14010–14021. doi: 10.1523/JNEUROSCI.1528-12.2012

Woods, D. L., Yund, E. W., Herron, T. J., and Ua Cruadhlaoich, M. A. I. (2010). Consonant identification in consonant-vowel-consonant syllables in speech-spectrum noise. *J. Acoust. Soc. Am.* 127, 1609–1623. doi: 10.1121/1.3293005

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Heslenfeld, D. J., and Rönnberg, J. (2012). Behavioural and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. *Brain. Lang.* 122, 103–113. doi: 10.1016/j.bandl.2012.05.006

Zion Golumbic, E. M., Poeppel, D., and Schroeder, C. E. (2012). Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain. Lang.* 122, 151–161. doi: 10.1016/j.bandl.2011.12.010

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 February 2014; accepted: 13 May 2014; published online: 02 June 2014.
*Citation: Moradi S, Lidestam B, Saremi A and Rönnberg J (2014) Gated auditory speech perception: effects of listening conditions and cognitive capacity. Front. Psychol. **5**:531. doi: 10.3389/fpsyg.2014.00531*
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology.
Copyright © 2014 Moradi, Lidestam, Saremi and Rönnberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 09 June 2015
doi: 10.3389/fpsyg.2015.00778

Relatively effortless listening promotes understanding and recall of medical instructions in older adults

Roberta M. DiDonato[1,2]* *and Aimée M. Surprenant*[1]

[1] Cognitive Aging and Memory Lab, Department of Psychology, Memorial University of Newfoundland, St. John's, NL, Canada, [2] Speech Language Pathology, Medicine Department, Eastern Health, St. John's, NL, Canada

Communication success under adverse conditions requires efficient and effective recruitment of both bottom-up (sensori-perceptual) and top-down (cognitive-linguistic) resources to decode the intended auditory-verbal message. Employing these limited capacity resources has been shown to vary across the lifespan, with evidence indicating that younger adults out-perform older adults for both comprehension and memory of the message. This study examined how sources of interference arising from the speaker (message spoken with conversational vs. clear speech technique), the listener (hearing-listening and cognitive-linguistic factors), and the environment (in competing speech babble noise vs. quiet) interact and influence learning and memory performance using more ecologically valid methods than has been done previously. The results suggest that when older adults listened to complex medical prescription instructions with "clear speech," (presented at audible levels through insertion earphones) their learning efficiency, immediate, and delayed memory performance improved relative to their performance when they listened with a normal conversational speech rate (presented at audible levels in sound field). This better learning and memory performance for clear speech listening was maintained even in the presence of speech babble noise. The finding that there was the largest learning-practice effect on 2nd trial performance in the conversational speech when the clear speech listening condition was first is suggestive of greater experience-dependent perceptual learning or adaptation to the speaker's speech and voice pattern in clear speech. This suggests that experience-dependent perceptual learning plays a role in facilitating the language processing and comprehension of a message and subsequent memory encoding.

Keywords: memory, hearing loss, aging, auditory processing, comprehension

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Helen Henshaw,
University of Nottingham, UK
Elaine Hoi Ning Ng,
Linköping University, Sweden

***Correspondence:**
Roberta M. DiDonato,
Cognitive Aging and Memory Lab, Department of Psychology, Memorial University of Newfoundland, 230 Elizabeth Ave., St. John's, NL A1B 3X9, Canada
rmd308@mun.ca

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *26 January 2015*
Accepted: *25 May 2015*
Published: *09 June 2015*

Citation:
DiDonato RM and Surprenant AM (2015) Relatively effortless listening promotes understanding and recall of medical instructions in older adults. Front. Psychol. 6:778. doi: 10.3389/fpsyg.2015.00778

Introduction

Adverse listening conditions that may hinder communication success arise from multiple sources. They may arise from within the speaker (imprecise articulation or accented speech), within the listener (hearing loss or cognitive-linguistic compromise) and/or within the environment (degraded transmission of the communication signal from telecommunication systems) (Mattys et al., 2009, 2012; Mattys and Wiget, 2011). By examining how speaker, listener and environmental sources of interference interact and influence language understanding and communication success, those factors or mechanisms that may also hinder or facilitate

learning and memory performance can be identified (McCoy et al., 2005). This could have many practical impacts. First, those components that are most amenable to intervention could be improved in order to affect functional performance of activities of daily living that require communication and memory of important instructions (IADLs). Second, understanding them will advance our knowledge of how age-related changes in sensory-perceptual abilities influence cognitive decline in the older adult and may provide opportunities for prevention.

The primary purpose of this study was to accomplish the following goals: (1) to examine whether a specific type of auditory enhancement, a message spoken with clear speech technique, relative to normal conversational speech results in better learning efficiency, immediate, and delayed memory performance (Bradlow et al., 2003); (2) to investigate whether a distractor (e.g., speech babble noise) decreases learning and memory performance similarly in both the conversational and clear speech listening conditions; and (3) to determine how individual differences in hearing-listening and cognitive-linguistic factors contribute to memory performance. Three sources that contribute to adverse listening conditions were examined: those that arise within the speaker (conversation vs. clear speech), the listener (hearing-listening or cognitive linguistic functioning) and the environment (noise vs. quiet). Further, due to the nature of the design, learning-practice effects were also considered in this study. Specifically it was important to determine if memory performance was influenced as a result of practice with the experimental tasks, specifically for the role of experience-dependent perceptional learning or adaptation to the speaker (Peelle and Wingfield, 2005).

A secondary purpose was to examine this in an ecologically valid manner that captures real-life listening, language comprehension, and memory performance that is pragmatically relevant for many older adults. One motivation to use ecologically valid methods and tasks is to generalize these findings to more typical communication scenarios that require dual-tasking such as learning a task while listening to instructions (Schaefer, 2014). Additionally, as Gilbert et al. (2014) suggested, enhanced speech intelligibility with ecologically valid methods is necessary for examining how speech perception and processing in more naturalistic communicative scenarios influences listening effort and memory in older adults. Another motivation is to address the criticism of cognitive-aging research that uses methods and tasks that are more relevant to university students and less relevant to older adults, particularly when comparing the groups' performance. The criticism is that the older adults' poorer performance could be attributed to reasons unrelated to cognitive-aging decline (older adults view tasks to be patently artificial and therefore are less motivated to perform) (Craik and Bialystok, 2006).

Age-related hearing loss (ARHL) can be defined as a combination of auditory perceptual and auditory processing deficits. These age-related changes in auditory perception and processing have been demonstrated to occur as early as middle age (e.g., 40–57 years old) (Working Group on Speech Understanding and Aging and the Committee on Hearing, Bioacoustics and Biomechanics (CHABA), 1988; Helfer and Vargo, 2009; Wambacq et al., 2009). The etiology of ARHL can be attributed to a combination of the auditory stressors that are acquired throughout the life span (e.g., trauma, noise, and otologic diseases) together with genetically controlled aging processes (CHABA, 1988). Older adults with clinically normal audiograms demonstrate less *dynamic* temporal processing abilities as compared to younger adults with normal hearing (Konkle et al., 1977; Gordon-Salant and Fitzgibbons, 1993). Additionally, a mixed-type hearing loss is also consistent with this definition of ARHL. Therefore, a broader definition of ARHL beyond the audiogram (high frequency sensori-neural hearing loss) was considered for this study, one that incorporates these other aspects of hearing-listening changes that interfere with signal processing for speech understanding (Anderson et al., 2011, 2012; John et al., 2012).

There is evidence that as we age, particularly around the 6th decade of life, our listening abilities are less precise and less efficient compared to younger adults in the 2nd to 3rd decades of life (CHABA, 1988). These age-related hearing-listening changes distort and degrade the stimuli (Rosen, 1992; Gordon-Salant and Fitzgibbons, 1993). These listening difficulties arise from at least three general areas: decreased audibility particularly in the high frequencies disrupting consonant discrimination (Humes, 2008), slowed temporal processing or adaptation (Peelle and Wingfield, 2005) interference with experience-dependent perceptual learning of the speaker's voice and speech pattern, and difficulty segmenting the target from a competing message (e.g., listening in noise). The listening-in-noise difficulty evident in the older adult arises from both domain-specific processes (such as auditory stream segregation) and domain-general cognitive-linguistic processes (such as attention, task switching, inhibition, and monitoring capacity) (Anderson et al., 2012, 2013; Humes et al., 2012; Amichetti et al., 2013).

Furthermore, several studies have shown that even mild hearing loss that has no measurable effect on speech understanding in quiet listening conditions can have substantial effects in noisy or other adverse conditions for both discriminating words (CHABA, 1988), and memory for words recognized (Rabbitt, 1990; Pichora-Fuller et al., 1995; Mattys et al., 2009, 2012; Ng et al., 2013).

The ability to understand spoken language is necessary for functional performance of instructional activities of daily living (IALDs) (e.g., use of medical instructions for medical adherence). Fundamental to comprehension and learning of an auditory-verbal message are sufficiently intact auditory perceptual-processing abilities and cognitive-linguistic functioning. These bottom-up (auditory perceptual-processing) and top-down (cognitive-linguistic) processes need to be efficiently recruited to effectively decode the message for communication success. Both implicit and explicit recruitment of these limited-capacity resources (Kahneman, 1973), perhaps as compensation (Bäckman and Dixon, 1992; Rönnberg et al., 2010; Wild et al., 2012) have been demonstrated to promote ease of language understanding in sub-optimal or adverse communication scenarios.

Rönnberg et al. (2008) used a working memory model for Ease of Language Understanding (ELU) to explain how perceptual

processes interact with cognitive processes for understanding. They proposed that it is the relative fidelity of the speech message that allows for the ease or automaticity of the match between the upstream sub-lexical features (phonology) and the target in the lexicon. Thus, when the fidelity is optimal, the match with the target occurs, at the exclusion of other competing targets in the lexicon, more rapidly and automatically due to implicit processes. When the fidelity of the message is low or suboptimal, the automatic matching processes of the sub-lexical features to the target in the lexicon is unsuccessful, resulting in a mismatch. The ELU model suggests that controlled processes are then required such that the sub-lexical, lexical, and semantic and conceptual representations from long-term memory are needed to further decode the speech signal. The match then occurs by way of explicit processes (Rönnberg et al., 2008, 2013). Thus, the re-allocation of explicit cognitive-linguistic resources for decoding of the speech signal results in fewer resources available for the learning and recall of the materials heard. Under optimal listening conditions fewer explicit resources are needed for comprehension, presumably because the perceptual features more closely match the listener's sub-lexical and lexical features in long-term memory. Optimizing the fidelity of the spoken message allows for more rapid and automatic-implicit perceptual learning of the speaker (Rudner et al., 2009) and more cognitive-linguistic resources will be available for comprehension, learning, and recall of the message (Wingfield et al., 1985, 1999, 2006; Wingfield and Ducharme, 1999).

One method to optimize the listening situation is to increase the fidelity of the speech message by using a style of speaking that increases the speech intelligibility. The "clear speech technique" is one in which the talker is instructed to produce the speech as if speaking to someone who is either hearing impaired or to one who is not a native speaker of the language (Ferguson and Kewley-Port, 2007). These were the instructions provided to the male speaker who produced the stimuli for our study. This "clear speech" technique resulted in an average speaking rate of 145 syllables per minute (spm). Relative to the original-conversational rate of the vignettes (192.5 spm), the clear speech rate was on the slower end of the normal speech rate (Goldman-Eisler, 1968); consistent with other studies that use this technique (Ferguson, 2012).

In addition to a slower rate of speech, other acoustic dimensions change by using the "clear speech" technique. The acoustic characteristics that give clear speech its intelligibility benefit are increased duration of vowels, longer and more frequent pauses, a larger consonant-vowel ratio, increased size of vowel space, decreased alveolar flapping, increased stop-plosive release, more variable voice fundamental frequency (F0), and greater variability in vocal intensity (Bradlow et al., 2003; Ferguson and Kewley-Port, 2007).

Although the use of clear speech has been demonstrated to enhance intelligibility of word and sentence discrimination in younger and older adults with and without hearing loss (Picheny et al., 1985; Ferguson, 2012) less is understood regarding its role for facilitating memory encoding. Gilbert et al. (2014) investigated intelligibility and recognition memory in noise for conversational and clear speech recorded in quiet and in response to the environmental noise (noise adapted speech-NAS) in young normal hearing adults. Results demonstrated that improved intelligibility for clear relative to conversational speech in noise improved recognition memory and that the NAS speech further enhanced intelligibility and recognition memory. Gilbert et al. (2014) concluded that naturalistic methods that simulate real-world communicative conditions for enhancing speech intelligibility have a role in improving speech recognition, comprehension, and memory performance in younger adults and may improve memory abilities for older adults.

Both sensory deficits (such as hearing loss) and cognitive impairments (such as memory difficulties) increase as a function of age and are highly correlated (Baltes and Lindenberger, 1997). In a comprehensive review of the literature, Schneider and Pichora-Fuller (2000) discussed a number of ways in which these sensory and cognitive declines could be related. They suggested that poor memory performance could be partially attributed to unclear and/or distorted perceptual information delivered to the cognitive/memory processes; the so-called "information-degradation hypothesis" (Schneider and Pichora-Fuller, 2000). In addition, several researchers (Rabbitt, 1968, 1990; Surprenant, 1999, 2007; Wingfield et al., 2005, 2006; Stewart and Wingfield, 2009; Tun et al., 2009; Baldwin and Ash, 2011) have argued that perceptual effort has an effect on cognitive resources with concomitant influences on memory performance. This is often referred to as the "effortfulness hypothesis."

According to the effortfulness hypothesis, if listening effort for decoding the verbal message comes at the cost of cognitive resources that would otherwise be shared with the secondary task of encoding information into memory, then decreasing listening effort should result in improved learning and memory performance. Further, those individuals with greater capacity in hearing-listening and cognitive–linguistic abilities would theoretically have more resources (Kahneman, 1973) to share between the two tasks (Rabbitt, 1968, 1990). Therefore, in order to determine how these bottom up and top down resources contributed to memory performance it was first necessary to examine the participant's unique abilities in hearing-listening and cognitive-linguistic functioning. Then, how these individual variables (hearing and cognition abilities) contribute to the memory performance by listening condition (conversational and clear) and by group (Quiet and Noise) can be examined.

In this study, we recruited older adults with a range from normal-to-moderately impaired hearing-listening abilities. They listened to medical instructions either in quiet or in the presence of background babble. Half of the sentences were presented in conversational speech and half in clear speech. The listeners were asked to repeat the stimuli as precisely as they could after each trial of listening. After a filled delay they were asked to recall all the information that they heard. We examined learning efficiency defined as the averaged amount of the stimuli repeated over the four trials to learn; immediate memory as the total of items repeated immediately; and the delayed memory as the total of items recalled after a delay period. We compared learning and memory performance within subjects for the two listening conditions (clear and conversational) and between subjects for the competition (quiet and noise). In addition, we measured the

individual's hearing-listening and cognitive-linguistic abilities to determine how these unique characteristics may have influenced the delayed memory performance in the two listening conditions for the two groups.

For theoretical and practical reasons, we examined how quickly the participant was able to learn the passages, how much they discriminated for immediate repetition and how much of the message they encoded for later free-recall. Theoretically, the question is whether these learning and memory processes in older adults are differentially affected by the change in listening condition. The intention is to identify the dissociable memory processing components that potentially contribute to a decline in memory for older adults (Salthouse, 2010).

Zacks et al. (2000) summarized the theoretical orientations in memory and aging and described three areas that differentiate the younger from the older adult; limited resources, processing speed, and inhibitory control.

Older adults are more limited in essential resources or self-initiated processing both at encoding and retrieval (Hasher and Zacks, 1979; Light, 1991; Craik et al., 1995). Relative to younger adults, older adults are more negatively affected by free-recall tasks, which require a higher degree of self-initiated processes. For the present study, the type of memory task chosen was free-recall. If the experimental manipulation to enhance the auditory stimuli improves the older adult's free-recall performance relative to conversational speech it will suggest that the age differences in free-recall, consistently reported by other authors (Salthouse, 2010), may be partially attributed to the effort in listening which consumes those same resources.

Older adults process information slower than younger adults (Park et al., 1996; Salthouse, 1996; Verhaeghen and Salthouse, 1997). According to Salthouse (1996) in situations in which time is restricted, the time required for the memory processes to rehearse or elaborately encode may be compromised by earlier processes, consuming the total time available to perform the task.

In relation to the present study, auditory enhancement (clear speech), which facilitates more timely and automatic processes for auditory perception and processing of the message, should free up time for those memory processes. In this way the auditory enhancements may facilitate faster perceptual learning or adaptation to the speaker's pattern. A larger learning effect (better learning or memory performance on 2nd trial of a task) indicates that the more automatic and timely auditory processing of the message for comprehension has allowed for more time available to rehearse or elaborately encode information for later recall. If learning effects differ by listening condition for the older adults, this finding suggests that some of the age-related slowing may be attributed to differences in perceptual learning of the speaker's pattern.

Older adults have less inhibitory control particularly for attention to the relevant contents of working memory. The increased mental clutter due to poorer inhibitory control increases the likelihood for sources of interference, both at encoding and retrieval (Hasher and Zacks, 1988; Zacks and Hasher, 1994, 1997; Hasher et al., 1999). In relation to the present study, the older adult with ARHL may experience an increase in mental clutter from the perceptual and lexical processing loads (Mattys and Scharenborg, 2014). Inhibiting this "noise" and maintaining attention to the task for both comprehension of the message and encoding into memory requires greater inhibitory control (or executive function) and working memory capacity for successful performance. In this way, the individual's executive control, working memory, and short-term memory is taxed more in adverse listening conditions relative to easier listening. Relevant to this study, those individuals with strengths in inhibitory control and working memory capacity should demonstrate better learning and memory performance, particularly for adverse listening conditions in which these resources are strained.

Both the ELU and the effortfulness hypotheses were considered for this study. According to the effortfulness hypothesis first described by Rabbitt (1968) and subsequently others (Tun et al., 2002, 2009; McCoy et al., 2005), while listening to typically spoken messages in degraded conditions, cognitive-linguistic resources are re-allocated for deciphering the message. This re-allocation of resources comes at the cost of those same resources for learning and memory encoding (Kahneman, 1973). The stimuli here were constructed in such a way as to optimize the auditory processing of the verbal message. The expectation is that the enhanced stimuli "clear speech technique" will mitigate those aspects of age-related hearing that interfere with communication success by reducing the perceptual, lexical, and cognitive loads (Mattys et al., 2012). In so doing, enhanced listening will free up those resources that are required for elaborate encoding for learning and remembering the passages.

Similarly, according to the ELU (Rönnberg et al., 2008), if the match between the stimuli and the long-term representation of the target in memory is automatic, then fewer explicit resources will be required for understanding the message. If we can enhance the clarity of the speech by using a style of speaking that promotes an intelligibility benefit, these same explicit cognitive-linguistic resources should become available for perceptual learning, comprehension, and elaborate encoding for later recall. Both of these hypotheses suggest that easier auditory processing of the message results in easier learning and recall. Also the suggestion is that resources for listening, learning, and remembering processes are limited and must be shared or re-allocated as needed (Gilbert et al., 2014).

If the hypotheses are confirmed, there should be a main effect of listening condition: Relative to conversational speech, enhanced listening will result in more efficient learning and better immediate and delayed memory performance. If the irrelevant speech-babble noise further interferes with processing of the targeted message then there will be a main effect of speech babble noise and an interaction of listening condition and group (Quiet vs. Noise). If found, the difference in memory performance between the two groups could be attributed to either energetic masking (Heinrich et al., 2008) of the stimuli, the noise covers up part of the sub-lexical acoustic information of the target; and/or a distractor effect, the noise distracts the listener's attention from the target (Lavie and DeFockert, 2003; Lavie, 2005; Mattys et al., 2009). In both scenarios, re-allocation of explicit cognitive-linguistic resources are required to "fill in" for what was missed to understand the message, while inhibiting the to-be-ignored

background and maintaining focus for processing of the ongoing message.

Materials and Methods

Participants

Ethics clearance was obtained from Memorial University's Interdisciplinary Committee on Ethics in Human Research (ICEHR) in accordance with the Tri-Council Policy Statement on Ethical Conduct involving Humans. Inclusion criteria: community dwelling-healthy older adults 55+ years old. Exclusion criteria: known medical events that may affect cognition (e.g., cardiovascular event, neurological event, or disease), failed cognitive screening, insufficient corrected vision for performing the experiment, and hearing loss that exceeded the capacity of the speakers (90 dBA). To determine the sample size required to detect a small effect size we used G*Power 3.1 (Faul et al., 2007) (Input: Effect size f = 0.26 α error probability = 0.05, Power (1-β error probability) = 0.95, Number of groups = 2, Number of measurements = 3, Correlations among repeated measures (learning efficiency, immediate, and delayed memory) = 0.5, Non-sphericity correction ε = 1. Output: Non-centrality parameter λ = 16.22, Critical F = 3.17, Numerator df = 2.0, Denominator df = 76.0). This suggested a total sample size of 40 participants. We over-recruited by 20% (e.g., 48 participants recruited) to account for attrition.

Forty-eight older adults were recruited to participate and were randomly assigned to either the Quiet (n = 24, 14 females) or Noise (n = 24, 12 females) group. This was accomplished by first generating a counterbalanced and randomized list for the two groups and the eight different orders for completing the experiment, then the participant was allocated to the pre-randomized group/order condition sequentially. Three participants wore hearing aids, two in the Quiet, and one in the Noise group. (See **Table 1** for demographic, hearing and cognitive characteristics means and standard deviations; see **Figure 1** for audiogram data.) Participants received $10 an hour for their participation.

TABLE 1 | Demographics, Hearing, and Cognitive Characteristics.

Characteristics	Quiet		Noise		Range
	M	*(SD)*	*M*	*(SD)*	*Min/Max*
DEMOGRAPHIC VARIABLES					
Age (years)	65.29	(6.16)	64.79	(6.94)	55/81
Education[a]	3.71	(1.04)	3.92	(1.06)	2/5
Health[b]	3.88	(0.74)	4.00	(0.83)	3/5
HEARING CHARACTERISTICS					
QuickSIN[c]	1.33	(1.39)	2.38	(1.64)	(–)1/(+)7*
HHIA Survey[d]	8.92	(12.62)	6.92	(10.27)	0/52
RPTA4 (dBHL)	16.04	(11.25)	20.99	(16.54)	(–)2.50/(+) 57.5
LPTA4 (dBHL)	19.90	(14.60)	20.05	(14.25)	(+)3.75/(+)65
Musicianship[e]	2.21	(2.67)	1.46	(2.13)	0/8
COGNITIVE CHARACTERISTICS					
FAS (words)[f]	43.04	(12.17)	42.63	(13.87)	17/73
BNT (words)[g]	56.79	(3.74)	55.00	(8.02)	23/60
Digits Back[h]	5.00	(0.93)	4.16	(1.30)	2/7*
L-Span (letters)[i]	18.13	(10.07)	17.04	(9.06)	0/42

Means and Standard Deviations.

[a]Education: self-reported category: 1, some High school; 2, High School; 3, some University/College; 4, University/college degree; 5, Graduate/professional degree.

[b]Health: self-reported category: 1, very poor; 2, poor; 3, good; 4, very good; 5, excellent.

[c]QuickSIN, Quick Speech-in-Noise measurement that provides a signal-to-noise ratio expressed as dB SNR loss, higher numbers indicate poorer abilities. Normal value, < +3 dB SNR loss (Killion, 2002).

[d]HHIA-Hearing Handicap Inventory for Adults: self-assessment; higher scores indicate greater perception of hearing handicap.

[e]Musicianship: interval scale 0–10 points (higher number reflects greater musicianship experience: 0, no music; 3, some previous music experience in past; 5, some past and current music; 10, full musician).

[f]FAS- verbal fluency-executive function task, higher number of words generated is better performance.

[g]BNT-Boston Naming Test, higher number of pictures correctly named is better performance.

[h]BackDigit Span-backwards digit span, mean number of digits reported for final 10 trials, higher number is better performance.

*[i]L-Span-Listening span, the sum total of letters recalled for each list length recalled with 100%. Larger number is better performance. *$p < 0.05$.*

Preliminary Measures

The purpose of these measures was to determine if an individual should be excluded from the study. No participant was excluded from the experiment based on the measures of vision, hearing, or the cognitive screening (e.g., passing score is >23) (Crum et al., 1993) the scores ranged from 27 to 30 on the Mini-Mental Status Examination (MMSE) (Folstein et al., 1975).

The following hearing-listening and cognitive-linguistic measures were obtained for all participants, the rationale for these measures and the standardized methods used are described in greater detail elsewhere (DiDonato, 2014).

Hearing-listening Measures

Audiometric tests were conducted in a single-walled sound attenuated chamber using a *Grason Stadler Instruments* Audiometer (GSI-61), Telephonics TDH50P headphones, E.a.r.Tone™ 3A insert earphones and free-field speakers calibrated to specification (American National Standards Institute ANSI, 2004). Standardized procedures with the TDH50P headphones were used to obtain pure-tone hearing thresholds for right (R) and left (L) ear. Pure tone average (PTA4) is the average of 0.5, 1, 2, and 4 kHz in dB HL (Katz, 1978). PTA4 was the metric used to indicate degree of auditory acuity deficit consistent with the WHO definition (PTA4 greater than 25 dB HL) (World Health Organization Prevention of Blindness and Deafness (PBD) Program, 2014). Speech Reception Threshold (SRT) is the threshold in dB at which one can repeat a closed set of words with 50% consistency (Newby, 1979). The Phonetically balanced (PB) max-most comfortable loudness level (PB max-MCL) is the intensity level measured in decibels in Hearing Level (dB HL), for which the participants achieved the highest accuracy for repeating phonetically-balanced (PB) word lists (Newby, 1979). The SRT and PB max-MCL were used to calculate the sensation level in which participants experienced the stimuli.

The Quick Speech-In-Noise test (QuickSIN): Etymotic Research, Elk Grove, IL; (Killion et al., 2004) is a standardized assessment of the ability to repeat/recall sentences from a target speaker (a female voice) in the presence of multi-talker babble

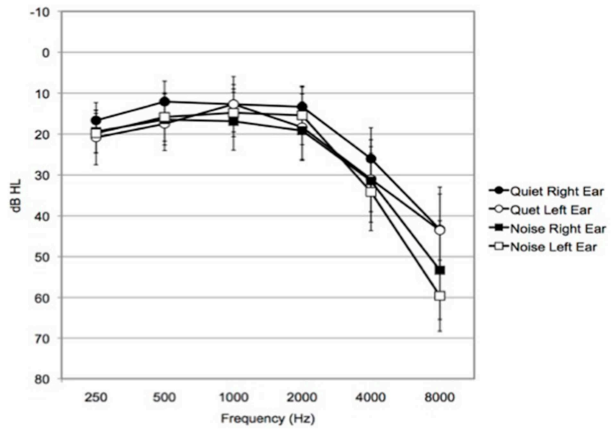


FIGURE 1 | Mean audiogram profile. Hearing thresholds of all participants in this study. Mean audiogram profile of Quiet group right ear and Quiet group left ear ($n = 24$), Noise group right ear and Noise group left ear ($n = 24$). Bars represent 95% confidence intervals.

at various levels of speech-in-noise ratios (SNRs). The target sentences were routed through the GSI-61 audiometer's external channel at 70 dB HL via the free-field speaker (Killion, 2002). The score is the signal-to-noise ratio (SNR), in decibels (dB), in which the listener recognizes the speech target correctly with 50% accuracy. A score of +7 dB SNR loss on the QuickSIN indicates that the individual needs the signal to be 7 dB louder than the competing speech noise in order to identify the sentences with 50% accuracy. Higher values reflect poorer listening-in-noise ability. The Hearing Handicap Inventory for Adults HHIA (Newman et al., 1991) is a standardized and normed self-assessment used clinically to determine the individual's self-perception of the degree to which they experience a handicap due to hearing loss (adapted from Hearing Handicap Inventory for the Elderly, HHIE (Ventry and Weinstein, 1982). The questions reflect both the social/situational and emotional consequences of hearing loss. The individual's response is yes (4 points), sometimes (2 points), or no (0 points). The score is the sum total of all the responses. A higher value reflects a greater perception of hearing handicap.

A musicianship score was calculated based on the responses to the demographic questionnaire regarding musical experience. The demographic questionnaire also included questions regarding age, education, occupation, health, medication use, and language(s) spoken (see Appendix A in supplementary Material). The musicianship classification score created for this study was an interval scale in which a higher value reflected more experience with music. Participants answered questions regarding exposure to music, age of onset of formal training, duration in years of musical performance, and the extent to which they were engaged in musical practice (e.g., hours/days per week). These questions were consistent with other studies that examine musical training and its relationship with auditory perceptual and processing abilities in behavioral and electrophysiological studies (Kraus and Chandrasekaran, 2010; Zendel and Alain, 2012, 2013). A composite score was calculated so that participants had a musicianship score from 0 to 10. A minimum score of 0 reflected no early music education, no formal lessons, and no instrumental or vocal performance presently or in the past. Maximum score of 10 reflected those who identify themselves as a musician (not necessarily professionally), started music education by 10 years of age or younger, had been musically active throughout their lifetime, had performed 12 years or greater, and those who currently perform on average at a minimum of 6 h weekly.

Cognitive-linguistic Measures

Listening span (L-span) is a working memory (WM) task that is similar to the reading span measure (Daneman and Carpenter, 1980). The rationale for using a WM span task in this study was that this type of span task is highly predictive for complex cognitive behaviors across domains such as understanding spoken language and reading comprehension (Just and Carpenter, 1992; St Clair-Thompson and Sykes, 2010). Participants heard a sentence and had to indicate whether the last word in the sentence was predictable or not predictable (mouse-click on the respective boxes on the computer screen). At the same time that they heard the sentence, they saw a letter on the computer screen. They were instructed to attend to the letters presented and after a series of sentences and letters, were cued to recreate the letter sequence in order. The sum total of all the list lengths, which were correctly recalled, is the score. Higher scores reflect better working memory. Backward digit span (Wechsler, 1981) is a task that correlates with other measures of cognitive function such as working memory capacity, but not so strongly that it measures the same construct (Conway et al., 2005; St Clair-Thompson, 2010). Participants heard lists of digits and recreated them in reverse order. The score reflects the mean number of digits recreated in reverse order for the final 10 trials. Boston Naming Test (BNT) is a subtest of the Boston Diagnostic Aphasia Examination (Kaplan et al., 2001). The BNT is a standardized and normed confrontation picture-naming task. Participants name 60 line drawings, 1 point for each correctly named item. The BNT has been found to have good internal consistency and high reliability (Goodglass et al., 2001). Verbal fluency measure (FAS) correlates with other metrics that measure executive function. Scores reflect the individual's cognitive flexibility, inhibition and response generation (Mueller and Dollaghan, 2013). Participants generate as many words as possible beginning with the letter "F," "A," and "S," given 1 min for each letter. The score is the total number of words generated.

Comparing Groups on Demographic, Hearing, and Cognitive Measures

There were no differences on demographic, hearing, and cognitive measures between the competition groups (Quiet/Noise) by ANOVA or Mann-Whitney *U*-tests (where appropriate) (smallest $p > 0.23$) except on the QuickSIN, $F_{(1, 47)} = 5.65$, $p = 0.02$, and Backward digit span, $F_{(1, 38)} = 5.36$, $p = 0.03$. The Quiet group demonstrated better listening-in-noise abilities, $M_{\text{Quiet}} = 1.33$ dB, $SD = 1.39$ dB, compared to the Noise group $M_{\text{Noise}} = 2.38$ dB, $SD = 1.64$ dB. The Quiet group demonstrated longer backward

digit span values ($M_{Quiet} = 5.00$, $SD = 0.93$), compared to the Noise group ($M_{Noise} = 4.16$, $SD = 1.30$). Due to an error in the program there were nine backward digits scores that had been incorrectly calculated (5 Quiet, 4 Noise); these values were not entered in the analysis for this measure. (**Table 1**).

There were unexpected *a priori* differences between the groups. If differences exist between the two competition groups for the learning and memory performance in the two listening conditions, these variables must be considered and understood in terms of their impact. The Quiet group's better listening-in-noise and short-term memory abilities could result in better learning and memory performance for the two listening conditions independent of the lack of noise (i.e., erroneously concluding that the noise interfered with performance). However, no main effect of group or interaction would suggest that these differences did not influence the result.

The Auditory-verbal Stimuli

Fictionalized medical prescription vignettes were created. The vignettes were thematic in nature and described the multiple steps needed to use specific medical prescriptions (see Appendix B in Supplementary Material for the two vignettes: medipatch and puffer-inhaler and training item). These vignettes were matched on many linguistic and non-linguistic aspects of speech to equate them as much as possible on the complexity of the stimuli, while at the same time maintaining their ecological validity (see **Table 2**). Both sets of prescription instructions comprised 10 sentences, with 37 critical units (CU) to report. The 37 CU were the content words within each phrase that carried the most important salient meaning for the practical purpose of using these fictional medications. Critical units may be a single word, compound word, or multiple words (e.g., breathe out, out of reach). The distribution of the CU throughout the vignette was arranged so that each third of the vignettes had similar numbers and distribution of items to recall. The two vignettes were spoken at their original-conversational rate, 192.5 (spm) and then these same vignettes were spoken using a slower hyper-articulated "clear speech" technique, (145 spm) (Baker and Bradlow, 2009).

The *clear speech* and the *conversational speech* vignettes in this experiment were subjected to acoustic analysis using Praat version 5.3.63 (Boersma and Weenink, 2014). Similar to Bradlow et al. (2003), total sentence duration, total number of pauses, average pause duration, F0 mean (Hz), F0 range (Hz), and the average vowel space range in F1 (mels) and F2 (mels) were examined. To calculate the vowel space in mels, the frequency (Hz) was converted to the perceptually motivated mel scale according to the equation by Fant (1973). Similar to Bradlow et al. (2003), when the speaker used a "clear speech" technique there was an increase in the overall duration, the number of pauses, a change in F0 mean and range, and increase in vowel space relative to when the conversational style speech technique was used. Thus, the clear speech vignettes reflect a temporal-spectral enhancement relative to the conversational speech vignettes (see **Table 3** for the characteristics of each vignette; **Figure 2** for Praat waveform). Avid Pro-tools 8.0.5 was used to manipulate the original sound files to ensure that the recordings were equated for loudness [root mean squared (RMS) amplitude] throughout the passages.

TABLE 2 | Linguistic aspects of the vignettes.

	Medipatch	Puffer
LINGUISTIC FEATURES		
Total words + (carrier)	100 (15)	89 (24)
Function words	25	26
Content words	75	63
#Syllables CU	73	73
Max #syllables in sentence	21	21
Min #syllables in sentence	3	4
Imperative phrases	11	12
Total #sentences (units)	10 (37)	10 (37)

Research Design

There was one between-subjects variable, competition (Quiet vs. Noise) and two within-subjects variables, listening condition (conversational vs. clear speech), and time of memory recall (immediate vs. delayed). This study used a modification of the learn-relearn paradigm (Keisler and Willingham, 2007). Participants listened to, immediately repeated what they had heard (immediate memory), and learned the vignettes as precisely as they could over a series of trials (learning efficiency). They then recalled the vignettes after the completion of 20 min of interference/filler tasks (delayed memory). The participants completed the study in two sessions on two separate days. In the first session they completed the vision screening, audiometric tests and the listening span (L-span). In the second session they completed the experiment as well as the other measures of hearing-listening and cognitive-linguistic abilities (included in the interference/filler task sets A and B).

Each participant listened to two passages (medipatch and puffer), one spoken with conversational and one in clear speech listening conditions, and all preliminary measures and filler/interference tasks (set A and set B). This resulted in eight different combinations of order conditions. The order in which participants performed the listening conditions, passages, or tasks (set A and B) was counterbalanced and participants were randomly assigned to one of the order conditions. An example of one of the orders is *EmA/DpB*. **Figure 3** illustrates the procedures for the second session, when the participant performed the experiment in two listening conditions. In this example, the participant experienced the relatively *Enhanced* listening condition first (clear speech through insertion ear phones) with the *medipatch* passage, completed the interference/filler *tasks set A*. At completion of the timer the participant then returned to the sound booth to recall the medipatch passage. There was a 5-min break (/) between the first and second listening condition. Then the participant experienced the second listening condition, the relatively *Degraded* listening condition (conversational speech through the speaker in sound field) with the *puffer-inhaler* passage, completed the interference/filler *task set B*. Again at

TABLE 3 | Acoustic Characteristics of Conversational (Conv.) and Clear speech.

Acoustic Measurement	Medipatch			Puffer		
	Conv.	Clear	Difference	Conv.	Clear	Difference
Avg. passage duration (s)	48.30	62.20	13.90	47.00	64.00	17.00
Total # of pauses	12.00	18.00	6.00	10.00	25.00	15.00
Avg. pause duration(ms)	7.60	7.50	0.10	7.10	9.50	2.40
F0 mean (Hz)	113.36	126.02	12.66	114.05	121.84	7.79
Vowel space F1 (mels)	748.02	775.53	27.51	655.32	690.49	35.17
F0 range (Hz)	233.25	308.00	74.74	317.22	356.39	39.17
Vowel space F2 (mels)	1368.41	1426.52	58.11	1442.81	1517.00	74.19

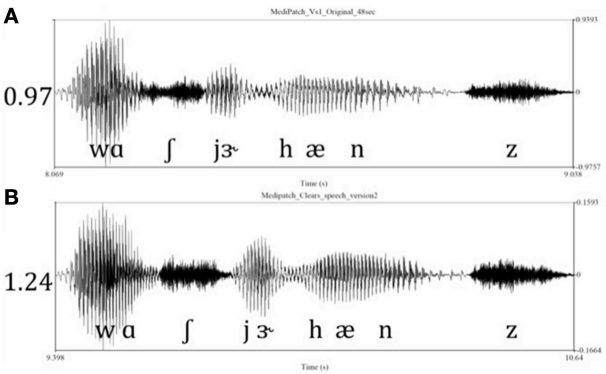


FIGURE 2 | The Praat waveforms: Two listening conditions. The waveforms depict the phrase "wash your hands" from the medipatch vignette. The two listening conditions: **(A)** 0.97 s, original format, conversational speech technique (196 spm); **(B)** 1.24 s, spoken with clear speech technique (152 spm). Note in clear speech, the temporal-spectral enhancement can be appreciated by the increased durations of the vowels and increased amplitudes of the waveform.

completion of the timer the participant returned to the sound booth to recall the puffer-inhaler passage.

Filler/interference tasks. The tasks had two purposes: (1) to provide a delay between listening and delayed recall and a filler activity; and (2) to assess participants on various cognitive and linguistic measures that were later used in the correlation analyses to examine the individual differences in relationship to memory performance. The tasks within each set were administered in the same order. Set A included the (FAS), the backward digit span task, the Philadelphia naming test items 1–87 (Roach et al., 1996), and a demographic questionnaire. Set B included the Philadelphia naming test items 88–175, the BNT, the MMSE, and the HHIA.

There were three dependent measures that were obtained for the two listening conditions as follows: *Learning efficiency* was operationally defined as the mean number of CU learned per trial, calculated using the total sum of the number of CU reported at each of the four trials of learning divided by the number of trials (4). In this way there was a single value for the learning efficiency during the conversational listening, and a single value for the learning efficiency during the clear condition. *Immediate memory* was operationally defined as the sum total of the CU that had been reported during any of the learning trials for that listening condition, to the maximum of a possible total of 37 units (e.g., 1st trial (15) reported CU, plus 2nd trial (5) new CU, plus 3rd trial (3) new CU, plus 4th trial (1) new additional units = 24 CU recalled immediately for that listening condition). *Delayed memory* was operationally defined as the total number of reported CU after the filler tasks for that listening condition, to the maximum of 37 CU.

Instructions

Participants were informed of the experimental tasks with a written script (see Appendix C in Supplementary Material) that was read aloud to them, while they read along. Answers to questions and redirections to the written instructions were provided prior to and during the training/practice item. They were instructed that they would have multiple trials (4) to learn each vignette and to repeat all that they had heard and remembered after each trial of listening. Participants were instructed that gist reporting was acceptable but were encouraged to use as close to verbatim as possible. The participants were not under any time constraint. Responses were spoken aloud and the responses were audio-recorded. Each trial of listening and then recall of the vignette was recorded into GarageBand '11 on a Macintosh computer for later transcription and off-line scoring. A single research assistant blinded to the listening condition/competition group coded the data.

A training item was created so that participants could understand the nature of the task with feedback provided during the training task, and to confirm that the intensity level determined during the audiometric testing as PB max-MCL was comfortably loud but not too loud. After completion of the training/practice the participant was reminded to perform the experiment as they had just done during the training.

Presentation of the Auditory Condition

The stimuli were routed from a MacBook Pro computer via Apogee One, a studio quality USB music interface, to the auxiliary channels of the GSI-61 to the transducers (insert earphones or free-field speaker). The intensity level was set at each individual participant's PB max-MCL obtained during the audiometric testing. This individualized audibility level is consistent with an intensity level that reflects their best performance for discriminating and repeating a list of open-set words in quiet in a sound attenuated chamber.

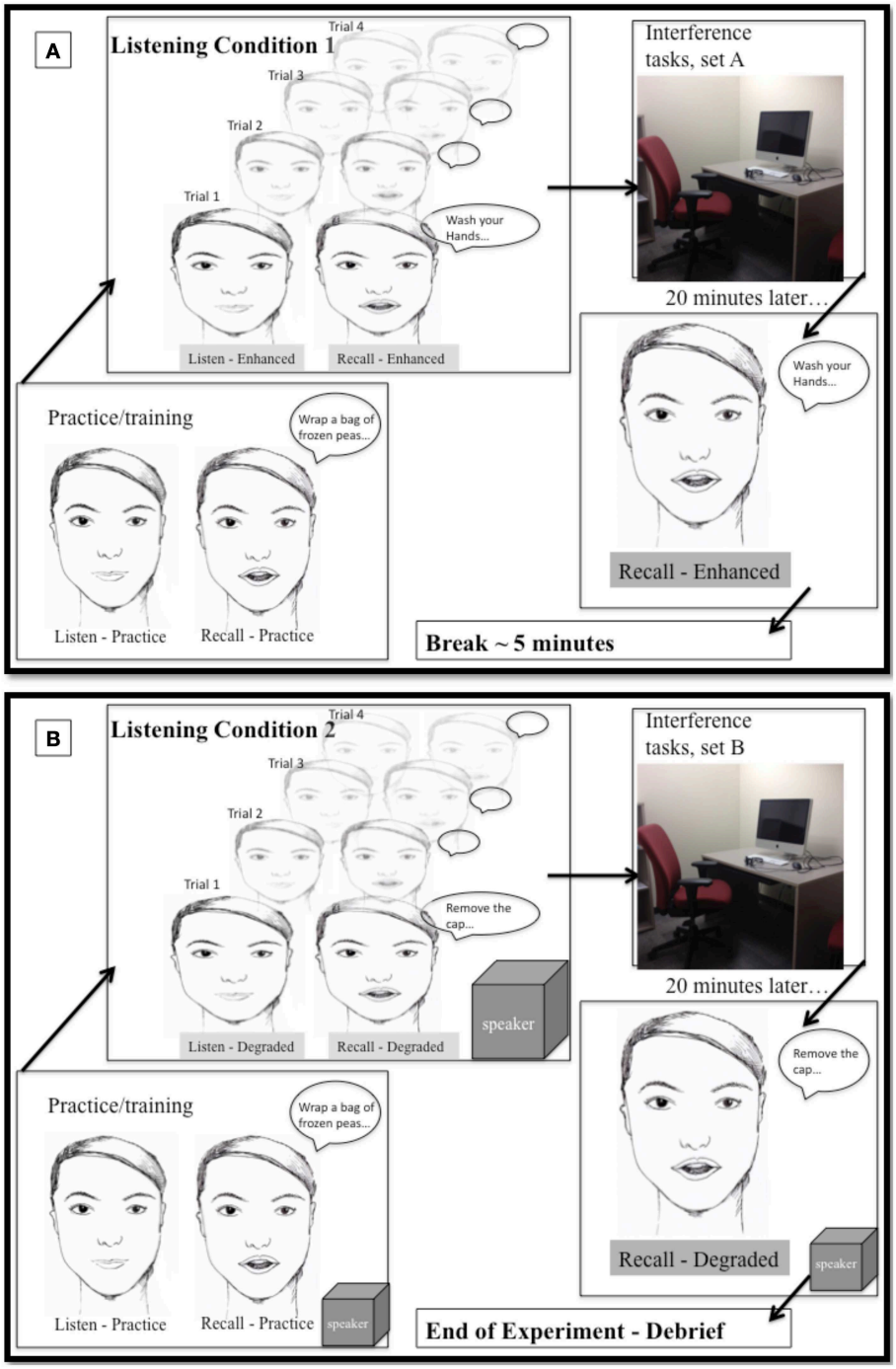


FIGURE 3 | Illustration of procedures for experiment (second session). (A) Top panel: Listening condition 1, enhanced (clear speech), via insertion earphones. Participant instructed and practice session. Trials of listening and recall × 4. Move to experiment room for 20 min of interference/filler tasks (set A). Move back to sound booth for delayed recall. Five min break. **(B)** Lower panel: Listening condition 2, degraded (conversational speech) via speaker. Participant re-instructed and practice session repeated. Trials of listening and recall × 4. Move to experiment room for 20 min of interference/filler tasks (set B). Move back to sound booth for delayed recall, end of experiment, debriefing.

Despite the advantage of using MCL in dB HL (see DiDonato, 2014), the actual sensation levels or hearing levels for the presentation of the stimuli may have varied by group. Therefore, the sensation level that the participants experienced was calculated for all participants in each group by subtracting the Speech Reception Threshold in dB from the MCL in dB HL, which indicates the sensation level in dB SL. There were no differences between the competition groups (Quiet/Noise) by

ANOVA for the sensation level presentation, $F_{(1, 47)} = 2.98$, $p = 0.09$ or for the MCL in dB HL, $F_{(1, 47)} = 0.96$, $p = 0.33$ (see **Table 4**).

Conversational Speech Listening Condition

The conversational speech was presented binaurally via a free-field speaker calibrated to a 1 kHz tone. Participants who wore hearing aids did so for this listening condition only. The free-field presentation was used for this listening condition to mimic listening in natural listening environments. All participants were seated and positioned 1 meter distance and 0 degree azimuth to the speaker. *The Noise group.* The conversational speech vignette and competing speech babble noise at +5 dB SNR were routed to the speaker. *The Quiet group*. The conversational speech vignette was routed to the speaker in quiet.

Clear Speech Listening Condition

The clear speech stimuli were presented binaurally via disposable 3A E.A.R.tone™ insert earphones. This was intended to further enhance listening by providing optimized signal-to-noise (SNR) benefit. This was done to simulate enhancements for listening by optimizing SNR benefit easily captured in the natural environment (i.e., heard with either a personal FM system, head phones, or through a looped hearing aid). The reality of an SNR benefit of the stimuli in Quiet with the insert earphones in an anechoic sound-attenuated chamber would be much less but perhaps not zero. Additionally, since the clear speech signal and the noise were transduced via the insert earphones simultaneously the SNR benefit would have been nullified for the Noise group. *The Noise group*. The clear speech vignette and competing speech babble noise at +5 dB SNR were presented simultaneously to the insert earphones binaurally. *The Quiet group*. The clear speech vignette was presented without speech babble noise to the insert earphones binaurally.

Results

To determine the consistency and accuracy of the coding of the participant sound files, one research assistant, blinded to the listening condition, coded all the participant files and then re-coded 21% of the total of the files randomly selected from the experiment. Intra-rater reliabilities for coding of blinded scoring were assessed using intra-class correlation coefficient (ICC) with a two-way mixed effects model and absolute agreement type (Shrout and Fleiss, 1979). The ICC for single measures for the reported-recalled CU for each trial was 0.98. An ICC value between 0.75 and 1.00 is considered excellent (Hallgren, 2012). The high ICC intra-rater reliabilities suggests that minimal amount of measurement error was introduced by the coding of the participants' sound files (Cicchetti, 1994).

TABLE 4 | Intensity level of stimuli presentation.

	Quiet		Noise	
	M	*(SD)*	*M*	*(SD)*
Sensation Level dB SL	45.63	(6.81)	41.88	(8.18)
MCL in dB HL	58.96	(6.08)	61.04	(8.47)

Order of Experiment Effects

There were eight different orders in which the participants completed the experiment. To determine whether the order of the experiment affected the participant's performance, a series of mixed design ANOVAs were conducted. The learning efficiency, immediate memory, and delayed memory scores were analyzed, with a 2 (listening condition: conversational vs. clear) × 2 (listen order: conversational first vs. clear first) × 2 (passage order: medipatch first vs. puffer first) × 2 (interference/filler task set order: Set A first vs. Set B first) mixed factors ANOVA, with listening condition as a within-subjects factor, and the three order variables as between-subjects factors. This was conducted for each of the dependent variables separately (see **Table 5** for all *F* and *p*-values).

Listening Condition Order and Listening Condition Interactions

There was an interaction between listening condition order (conversational-clear vs. clear-conversational) and listening condition on learning efficiency, $F_{(1, 40)} = 10.68$, $p = 0.002$, on immediate memory, $F_{(1, 40)} = 5.91$, $p = 0.02$, and on delayed memory, $F_{(1, 40)} = 4.04$, $p = 0.05$. This interaction is as follows: Performance was always better for the subgroups who experienced the listening condition as their second listening task compared to the subgroups who experienced that same listening condition as their first listening task (**Figure 4**).

Learning efficiency was better for second vs. first listening condition in both the conversational listening condition, $M_{\text{first-conversational}} = 19.66$, $SD = 5.81$, $M_{\text{second-conversational}} = 21.94$, $SD = 5.40$; and the clear listening condition, $M_{\text{first-clear}} = 21.03$, $SD = 6.75$, $M_{\text{second-clear}} = 23.09$, $SD = 5.73$.

Immediate memory performance was better for second vs. first listening condition in the conversational listening condition, $M_{\text{first-conversational}} = 28.79$, $SD = 5.38$, $M_{\text{second-conversational}} = 30.42$, $SD = 4.51$; and the clear listening condition, $M_{\text{first-clear}} = 29.63$, $SD = 5.79$, $M_{\text{second-clear}} = 31.33$, $SD = 4.43$.

Delayed memory was better for second vs. first listening condition in the conversational listening condition, $M_{\text{first-conversational}} = 22.83$, $SD = 5.85$, $M_{\text{second-conversational}} = 25.08$, $SD = 6.01$; and the clear listening condition, $M_{\text{first-clear}} = 24.54$, $SD = 6.73$, $M_{\text{second-clear}} = 25.21$, $SD = 6.38$.

This reflects general learning-practice effects, which were greater for the conversational (heard clear first) compared to the clear (heard conversational first) condition.

Post-hoc paired samples *t*-test (Bonferroni correction, alpha = 0.025) revealed that listening-order influenced the dependent variables differentially for the listening conditions. Conversational-1st order resulted in a significant difference in the two speech styles; for learning efficiency, $t_{(23)} = 3.60$, $p = 0.002$; immediate memory, $t_{(23)} = 2.49$, $p = 0.021$; and marginally significant for delayed memory, $t_{(23)} = 1.90$, $p = 0.07$. However, clear-1st order resulted in no difference in performance for listening conditions for the dependent

TABLE 5 | Order of Experimental Effects and Interactions.

Variables	$F_{(1,40)}$	P
LEARN EFFICIENCY		
Listening Condition	3.63	0.06
Listening Condition*Listening Order	10.68	***0.002**
Listening Condition*Passage Order	0.14	0.72
Listening Condition*Interference Order	1.33	0.26
Listening Condition*Passage Order*Listen Order	0.87	0.36
Listening Condition*Passage Order*Interference Order	3.05	0.09
Listening Condition*Listen Order*Interference Order	0.31	0.58
Listening Condition*Passage Order*Listen Order*Interference Order	0.10	0.75
IMMEDIATE MEMORY		
Listening Condition	1.63	0.21
Listening Condition*Listening Order	5.91	***0.02**
Listening Condition*Passage Order	2.13	0.15
Listening Condition*Interference Order	0.09	0.76
Listening Condition*Passage Order*Listen Order	0.63	0.43
Listening Condition*Passage Order*Interference Order	5.91	***0.02**
Listening Condition*Listen Order*Interference Order	0.02	0.90
Listening Condition*Passage Order*Listen Order*Interference Order	0.00	0.95
DELAYED MEMORY		
Listening Condition	1.60	0.21
Listening Condition*Listening Order	4.04	***0.05**
Listening Condition*Passage Order	0.05	0.82
Listening Condition*Interference Order	3.59	0.07
Listening Condition*Passage Order*Listen Order	0.40	0.53
Listening Condition*Passage Order*Interference Order	0.16	0.69
Listening Condition*Listen Order*Interference Order	0.00	1.00
Listening Condition*Passage Order*Listen Order*Interference Order	0.21	0.65

p-value bolded denotes significant.

variables, (all values for $t < 1$, $p > 0.34$). For example, when comparing the within-subject differences between the two speech styles (conversational vs. clear), there is a much smaller and non-significant differences when clear speech is heard first, where the difference between the two speech styles are significantly greater when conversational speech is heard first. **Figure 4** illustrates this difference for Delayed memory performance, gray bars represent the subgroup Clear second (25.21) − Conversational 1st (22.8) = 2.41; compared to the white bars, the subgroup Clear first (25.54) − Conversational 2nd (25.08) = 0.54. This larger and significant difference between the within-subject variable (conversational vs. clear listening condition) for the Conversational-1st is evident in both learning

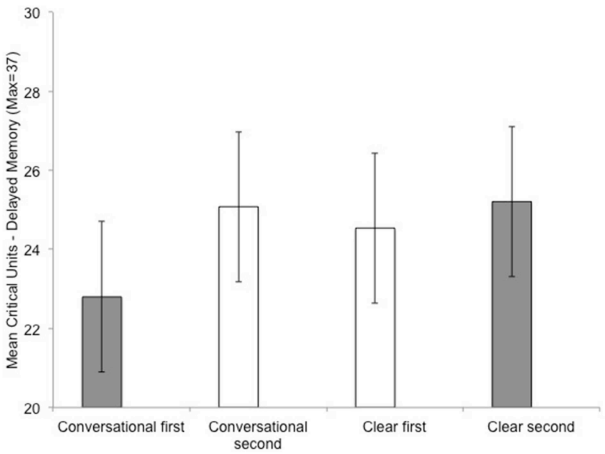


FIGURE 4 | Comparing learning effects with listening condition (conversational and clear speech) on delayed memory performance for the Quiet/Noise groups combined. First/second indicates the order in which the participant performed that experimental listening condition. The color of the bars differentiates the between-subject listening order in which they experienced the listening condition: Gray bars represent the subgroup of participants who listened in conversation first/clear second; white bars represent the subgroup of participants who listened in clear first/conversation second. Error bars are the standard error of the mean.

efficiency performance, 3.43 units, compared to Clear 1st a non-significant difference of 0.91; as well for the immediate memory performance, Conversational-1st, 2.54 units, compared to Clear 1st a non-significant difference of 0.79.

As a result of these interactions between listening-order and listening condition, listening order was entered as a covariate for further hypothesis testing of learning efficiency, immediate, and delayed memory performance between the Quiet and Noise groups in the conversational and clear listening conditions.

Passage, Interference/filler task, and Listening Condition Interactions

There was no effect of order or interactions for passage (e.g., medipatch vs. puffer) or interference/filler task set on Learning efficiency or Delayed memory performance (see **Table 5** for F and p-values). However, there was a 3-way interaction among passage (medipatch-puffer), interference/filler task (set A or B), and listening condition on immediate memory performance, $F_{(1, 40)} = 5.91$, $p = 0.02$.

The three-way interaction indicated that for the conversational speech listening conditions, those in the puffer passage with the interference task set A, immediately recalled more units, $M_{\text{conversational/puffer-set A}} = 32.75$, $SD = 3.47$, than the other 3 passage × interference task combinations, $M_{\text{conversational/puffer-set B}} = 28.50$, $SD = 5.33$, $M_{\text{conversational/medi-set A}} = 27.67$, $SD = 5.69$, $M_{\text{conversational/medi-set B}} = 29.50$, $SD = 4.10$; this was not the case in clear speech listening, the four subgroups are more similar, $M_{\text{clear/puffer-set A}} = 31.17$, $SD = 4.11$, $M_{\text{clear/puffer-set B}} = 29.83$, $SD = 6.42$, $M_{\text{clear/medi-set A}} = 31.42$, $SD = 5.11$, $M_{\text{clear/medi-set B}} = 29.50$, $SD = 5.21$.

As a result of the interactions noted above, listening condition order, passage order, and interference task order, were entered as covariates for further hypothesis testing for the differences of immediate memory between the groups (Quiet and Noise) in the conversational and clear listening conditions.

Listening Condition, Competition, and Interaction Effects on Learning and Memory Performance

Learning efficiency, immediate memory and delayed memory scores were analyzed with a 2 (competition: Quiet, Noise) × 2 (listening condition: conversation, clear speech) mixed design ANOVA in which listening condition was entered as the repeated measure within-subject variable and competition was a between-subject variable.

Effects of Listening Condition for Learning Efficiency, Immediate and Delayed Memory

There were main effects of listening condition on learning efficiency, $F_{(1, 45)} = 13.48$, $p = 0.001$, on immediate memory, $F_{(1, 43)} = 6.35$, $p = 0.02$, and on delayed memory, $F_{(1, 45)} = 5.51$, $p = 0.02$. The clear speech listening enhancements improved learning efficiency on average by 1.26 CU learned per trial and improved immediate and delayed recall on average by approximately 1 critical unit (see **Table 6**).

Effect of the Competition: Speech Babble Noise vs. Quiet

There were no main effects of the between-subject variable (competition: noise vs. quiet) on learning efficiency, immediate memory or delayed memory (all values for $F < 1, p > 0.57$).

Interaction Effects of Listening Condition and Competition

There were no significant interactions of listening condition by competition for learning efficiency, immediate memory or delayed memory (all values for $F < 1$, $p > 0.33$). The Quiet and the Noise groups were similarly affected by the "clear" speech enhancement to the listening condition.

TABLE 6 | Quiet and Noise groups for Learning Efficiency, Immediate, and Delayed Memory performance in conversational and clear listening conditions.

Dependent variable	Quiet		Noise		Total	
	M	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
LEARNING EFFICIENCY						
Conversation	20.99	(6.34)	20.60	(5.03)	20.80	(5.67)
Clear	22.15	(5.92)	21.98	(6.75)	22.06	(6.28)
IMMEDIATE MEMORY						
Conversation	29.88	(5.15)	29.33	(4.90)	29.60	(4.90)
Clear	30.92	(5.59)	30.04	(4.80)	30.48	(5.17)
DELAYED MEMORY						
Conversation	24.79	(6.81)	23.13	(5.01)	23.96	(5.97)
Clear	25.00	(6.25)	24.75	(6.87)	24.88	(6.50)

Means and Standard Deviations (CU).

Delayed Memory Performance and the Relationship with Hearing-listening and Cognitive-Linguistic Abilities

Correlation analyses were conducted to further explore the unique contribution of the individual's hearing-listening and cognitive-linguistic abilities on delayed memory performance in the conversational and clear speech listening conditions for the two groups (Quiet and Noise) separately. The rationale to conduct this analysis for only the delayed memory performance variable was based on the following. First, all three dependent variables showed similar patterns: the clear speech technique relative to the conversational listening condition resulted in better performance for learning efficiency, immediate, and delayed memory performances (approximately one additional critical unit reported). Second, these dependent variables were significantly and highly correlated with each other (see **Table 7** for correlation matrix of the dependent variables). Finally, important for the ecological validity of this study, the delayed memory variable was the metric that would support functional memory performance relevant to medical adherence.

The variables that reflected the hearing-listening ability as it relates to ARHL included in this analysis were LPTA4 and RPTA4, QuickSIN scores, the Hearing Handicap Inventory for Adults (HHIA), and musicianship score. The variables that reflected the cognitive-linguistic characteristics included in this analysis were as follows: auditory working memory as measured by L-span, executive function measured by verbal fluency task (FAS), lexical ability as measured by the word retrieval-picture naming task (BNT), and immediate memory as measured by the backwards digit span (Digits Back). The memory measures that were included in these correlation analyses were the delayed memory performance in the conversational and in the clear listening condition. These relationships were examined separately for the Quiet and the Noise groups.

Hearing-listening Abilities and Delayed Memory Performance

There were no correlations for LPTA4 and RPTA4; HHIA, QuickSIN, and Musicianship scores with delayed memory in the conversational and clear listening conditions in either the Quiet group or the Noise group when these groups are examined separately (see **Tables 8**, **9**, **10**).

TABLE 7 | Correlations between dependent variables for conversational (conv.) and clear listening.

	Delayed Conv.	Delayed Clear	Immediate Conv.	Immediate Clear	Learn Conv.
Delayed Clear	0.67**				
Immediate Conv.	0.84**	0.68**			
Immediate Clear	0.52**	0.84**	0.49**		
Learn Conv.	0.84**	0.73**	0.87**	0.58**	
Learn Clear	0.62**	0.93**	0.57**	0.91**	0.65**

$^{**}p < 0.01$.

TABLE 8 | Correlation analysis between delayed memory performance in the conversational (conv.) and clear listening conditions and hearing and cognitive abilities–Both groups.

	Delay Conv.	Delay Clear	LPTA4	RPTA4	HHIA	Quick SIN	Musician	L-Span	Digits Back	FAS
Delay Clear	0.67**									
LPTA4	−0.17	−0.17								
RPTA4	0.06	−0.13	0.63**							
HHIA	−0.06	−0.07	0.56**	0.32*						
Quick SIN	−0.24	−0.27	0.23	0.24	0.18					
Musician	0.17	0.20	−0.02	0.05	−0.14	−0.45**				
L-Span	0.39**	0.28	−0.24	−0.26	−0.19	−0.26	0.28			
Digits Back	0.48**	0.47**	−0.12	−0.04	0.08	−0.48**	0.25	0.43**		
FAS	0.53**	0.44**	−0.17	−0.32**	−0.15	−0.27	0.20	0.52**	0.29	
BNT	0.56**	0.55**	−0.04	−0.05	0.04	−0.34*	0.18	0.30*	0.30	0.36*

*$p < 0.05$, **$p < 0.01$.

TABLE 9 | Correlation analysis between delayed memory performance in the conversational (conv.) and clear listening conditions and hearing and cognitive abilities–Quiet group.

	Delay Conv.	Delay Clear	LPTA4	RPTA4	HHIA	Quick SIN	Musician	L-Span	Digits Back	FAS
Delay Clear	0.61**									
LPTA4	−0.25	−0.14								
RPTA4	−0.07	−0.28	0.57**							
HHIA	−0.08	−0.08	0.75**	0.20						
Quick SIN	−0.09	−0.18	0.07	0.17	0.16					
Musician	0.05	0.22	0.13	0.27	−0.11	−0.50*				
L-Span	0.36	0.28	−0.23	−0.28	−0.13	−0.45*	0.44*			
Digits Back	0.44	0.20	0.19	0.27	0.04	−0.16	0.43	0.30		
FAS	0.63**	0.43*	−0.10	−0.19	−0.05	−0.31	0.09	0.43*	0.47*	
BNT	0.64**	0.77**	−0.17	−0.33	0.10	−0.29	0.21	0.39	0.54*	0.51*

*$p < 0.05$, **$p < 0.01$.

TABLE 10 | Correlation analysis between delayed memory performance in the conversational (conv.) and clear listening conditions and hearing and cognitive abilities–Noise group.

	Delay Conv.	Delay Clear	LPTA4	RPTA4	HHIA	Quick SIN	Musician	L-Span	Digits Back	FAS
Delay Clear	0.78**									
LPTA4	−0.08	−0.19								
RPTA4	0.24	−0.03	0.72**							
HHIA	−0.06	−0.06	0.33	0.49*						
Quick SIN	−0.34	−0.35	0.39	0.22	0.29					
Musician	0.33	0.18	−0.21	−0.08	−0.22	−0.36				
L-Span	0.44*	0.27	−0.25	−0.25	−0.28	−0.07	0.04			
Digits Back	0.49*	0.59**	−0.34	−0.05	0.10	−0.51*	0.01	0.44		
FAS	0.46*	0.44*	−0.23	−0.40	−0.27	−0.26	0.33	0.61	0.13	
BNT	0.62**	0.50*	0.03	0.08	−0.002	−0.33	0.16	0.28	0.21	0.33

*$p < 0.05$, **$p < 0.01$.

However, when the entire sample was analyzed there were significant correlations with LPTA4, $r = 0.56$, $p < 0.001$; and with RPTA4, $r = 0.32$, $p = 0.03$ and self-perception of hearing handicap (HHIA); and a significant positive correlation of musicianship and listening-in-noise ability, (QuickSIN), $r = -0.45$, $p = 0.001$. Higher musicianship scores correlated with lower QuickSIN scores or better listening-in noise abilities. This is consistent with studies that examine the relationship of degree of musicianship and perception of speech-in-noise (Parbery-Clark et al., 2009, 2012). Those with more musical

training, for longer periods of time, starting at a younger age, demonstrate superior temporal processing, which supports better listening-in-noise abilities (Kraus and Chandrasekaran, 2010; Zendel and Alain, 2013). When considering the operationalized values of effect size as recommended by Cohen (1992), in which correlations >0.1 are considered small, >0.3 are considered medium, and >0.5 are considered large effect sizes. The above significant values ranged from medium to large effect sizes.

Although these hearing-listening abilities were not significantly related to delayed memory for the two listening conditions, generally the direction of the weak relationship of ARHL and memory performance was in the expected negative direction. As well, the hearing-listening measures did correlate with each other in the expected ways. For example, there were large effect sizes for the relationship between left and right acuity deficits and perception of hearing handicap (Newman et al., 1991), and a medium-large effect size of the relationship of musicianship and listening-in-noise abilities.

Cognitive-linguistic Abilities and Delayed Memory Performance

L-span: working memory ability and delayed memory performance

There was a significant positive correlation for the L-span scores and delayed memory for the Noise group in the conversational, $r = 0.44$, $p = 0.03$, but not in the clear, $r = 0.27$, $p = 0.20$, listening condition. There were no significant correlations for the L-span scores and delayed memory performance for the Quiet group for the conversational, $r = 0.36$, $p = 0.08$, and for the clear, $r = 0.28$, $p = 0.18$ listening condition. The magnitude of the effect decreased when the listening condition was more favorable as in the clear speech without the competing noise, in which it became non-significant.

Backward digit spans: short-term memory ability and delayed memory performance

In view of the fact that there were missing backward digit span scores, which most likely reflected poorer values, these results should be considered with some caution. There were significant positive correlations for the backward digit span scores and delayed memory for the Noise group in the conversational, $r = 0.49$, $p = 0.03$, and for the clear, $r = 0.59$, $p = 0.006$, listening condition. There were no significant correlations for the backward digit span scores and delayed memory performance for the Quiet group for either the conversational, $r = 0.44$, $p = 0.06$, or the clear, $r = 0.20$, $p = 0.41$, listening conditions.

When the entire sample was examined, there were significant positive correlations between backward digits spans and memory performance for both the conversational, $r = 0.49$, $p = 0.002$, and the clear, $r = 0.47$, $p = 0.003$, listening conditions. The magnitude of the effect became smaller when the listening condition was more favorable as in the clear listening or without competing noise.

FAS: Executive function ability and delayed memory performance

There were positive correlations of the FAS scores and delayed memory for the Noise group in the conversational, $r = 0.46$, $p = 0.02$, and for the clear, $r = 0.44$, $p = 0.03$, listening condition. There were positive correlations of the FAS scores and delayed memory for the Quiet group in the conversational, $r = 0.63$, $p = 0.001$, and the clear listening, $r = 0.43$, $p = 0.04$. The magnitude of the effect became smaller when the listening condition was more favorable in the clear speech listening condition. However, it is interesting to note that the magnitude of the relationship of executive function and delayed memory was the greatest in the Quiet group in the conversational listening condition, which is an unexpected finding that will be considered in more detail below.

Boston Naming Test (BNT): Lexical ability (naming/verbal fluency) and delayed memory performance

There were positive correlations for the BNT scores and delayed memory for the Noise group in the conversational, $r = 0.62$, $p = 0.001$, and the clear, $r = 0.50$, p = 0.01, listening condition. There were correlations for the BNT scores and delayed memory for the Quiet group in the conversational, $r = 0.64$, $p = 0.001$, and the clear, $r = 0.77$, $p < 0.001$, listening condition. The magnitude of the effect became *greater* when the listening condition was most favorable, that is in the clear speech listening condition without competing noise.

Summary of cognitive-linguistic abilities and delayed memory performance in the conversational and clear listening for the Quiet and Noise groups

When the entire sample was analyzed, as well as when the two groups (Quiet and Noise) were analyzed separately, there were medium to large effects of the cognitive-linguistic measures on delayed memory for the conversational and clear speech listening conditions. The magnitude of these effects generally became smaller when the listening condition was more favorable as in the Quiet group or in the clear speech enhancement (**Tables 8–10**).

Discussion

The purpose of this study was to examine how auditory perception and processing of a relatively enhanced speech message (clear vs. conversational speech) affected perceptual learning efficiency, immediate, and delayed memory performance in older adults with varying levels of hearing-listening abilities. This was examined with ecologically valid methods to assess how the older adult's learning and memory performance is influenced based on real-life listening scenarios, with relevant materials and with enhancements that could be reasonably achieved.

Ultimately the research question proposed was whether ease of perceptual processing (ELU hypothesis Rönnberg et al., 2008) or effortless listening (effortfulness hypothesis, Rabbitt, 1968) mitigates the distortions from ARHL in quiet and noisy listening and promotes better learning and memory. The clear speech relative to conversational speech in this study promoted

intelligibility similar to other studies that examined speech perception in younger and older adults (Ferguson, 2012). The slower rate, increased pauses, and acoustic changes (increased vowel space, F0 mean and range) enhanced the temporal-spectral aspects of the stimuli such that it was more similar to how the younger adult perceives speech compared to how the older adult typically perceives speech. Relative to younger adults with normal hearing, older adults with normal audiograms have been found to demonstrate less stable and less precise temporal processing of specific speech cues such as timing, frequency, and harmonics which interferes with speech discrimination (Anderson et al., 2012). These auditory temporal-spectral processes are necessary for discrimination of phonemes, morphemes and the regularities in the speaker's voice and speech pattern (Rosen, 1992). The stability of the acoustic information allows one to detect the regularities of the input over time. Optimal auditory perceptual ability allows one to temporally process and perceptually learn and adapt to the variability of the speaker, even within a single conversation (Mattys et al., 2012). The speech was optimized in this way to provide the older adult with the psycho-acoustic perception of speech more similar to how the younger adult experiences the stimuli (audible, slower, more distinctive).

The expectation was that clear speech would ease or decrease the effort for the experience-dependent perceptual learning of the auditory-verbal message, such that the older adult can adapt to the speaker's speech and voice pattern more efficiently, and stay attendant to the linguistic processing of the targeted message. As Salthouse (2010) states, "the most convincing evidence that the causes of a phenomenon are understood are results establishing that the phenomenon can be manipulated through interventions" (p. 157). Indeed this was the intent of the current study. Since learning and memory performance improved due to the behavioral intervention (listening enhancements) that manipulated those specific factors that were theoretically hypothesized to cause the phenomenon of poorer learning/memory performance, then these results support the hypothesis.

There are both theoretical and practical implications of these findings. Broadly defined, ARHL in older adults may indeed be contributing to age-related cognitive memory decline. Optimizing listening scenarios may significantly influence the functional performance of the older adult for IADLs.

Strengths in cognitive-linguistic abilities were positively associated with delayed memory performance with the magnitude of this effect greater in the relatively adverse listening (conversational speech). Larger effect sizes for cognitive-linguistic abilities on delayed memory performance in conversational vs. clear speech in a within-subject design suggests that indeed fewer explicit cognitive resources were required for deciphering the message in the enhanced listening.

These results are consistent with both the ELU and the effortfulness hypotheses in that making the speech audible and clearer enhanced learning and memory performance in older adults. Thus, the results of this study shed light on how sensory perception and processing declines in the older adult affect the implicit experience-dependent perceptual learning processes. This disruption to the perceptual learning processes then has cascading effects on higher-level cognitive-memory processes, delayed memory performance.

Learning-practice Effects: Order of Listening Condition and Delayed Memory Performance

The significant interactions between the order of the presentation of the listening condition (conversational-clear vs. clear-conversational) and listening condition on learning efficiency, immediate and delayed memory performance in this study, are consistent with the extant literature describing a *learning-practice effect* and the related *learning curve*. A practice or learning effect is described as more positive scores (e.g., faster, more accurate, higher consistency, more efficient) with experience of task over subsequent trials of the same type of task or test. This learning-practice effect and the classic s-shaped learning curve (progress plotted on the y axis as a function of time/trials on the x axis) has been described to occur on the simplest perceptual-motor tasks as well as complex cognitive tasks (Ritter and Schooler, 2001). It is evident in educational testing, clinical neuropsychological tests, and in research with test-retest experimental designs (Hausknect et al., 2006). Learning effects may be affected by familiarity with task, decreased anxiety with repeated trials, and employment of strategies learned and transferred to the subsequent trials (Ritter et al., 2004).

The design and methods employed in this study were conducted in such a way that these learning-practice effects were anticipated (participants randomly assigned to the counterbalanced order of the variables), investigated (order effects examined); and controlled for in the analyses (entered listening-order as covariate).

Learning-practice Effect Benefit on Delayed Memory Performance

Pure listening condition effects (i.e., without learning-practice effects) can be appreciated by examining the subgroups' (N = 24) first listening conditions (conversation first vs. clear first). Delayed memory performance is similarly improved in clear vs. conversation in quiet (+1.5 units) and noise (+1.92 units). This supports the statistical finding of the clear speech enhancement improving delayed memory performance in quiet and noise conditions. (**Figure 4**).

A learning effect benefit is defined as previous experience with the task or test improving performance compared to no previous experience. It is quantified as the difference in delayed memory performance between the subgroups who had that listening condition as their second condition and the subgroups who had that same listening condition first (i.e., no prior experience with doing the experiment). For example, for delayed recall Clear 2nd − Clear 1st = +0.67; Conversational 2nd − Conversational 1st = +2.25. The reported interaction is that the learning effect benefit is differentially influenced by which listening condition was first. The benefit of experiencing the experiment first with conversational speech only increased the clear speech performance over the "pure listening condition effect" by +0.67. Where the benefit of experiencing the experiment first with clear speech increased the conversational speech performance over the "pure listening condition effect" by +2.25. In this way,

conversational speech listening as the first listening condition provided less of a learning-practice effect benefit.

The learning-practice effect may be attributable to the fact that this subgroup of participants who had the second listening task as the conversational speech listening condition had the benefit of learning how to do the task first in their first listening condition (i.e., clear listening condition). They were able to *perceptually learn and adapt to* the speaker's voice and speech characteristics more easily after that first clear listening condition. Further, the finding that the magnitude of the relationship of executive function and delayed memory performance was the greatest in the Quiet group in the conversational listening condition indicates that strengths in this cognitive ability contributed to successful performance perhaps as compensation (Bäckman and Dixon, 1992; Wild et al., 2012).

These results suggest the following: (1) The "clear" speech relative to conversational speech promotes an additional perceptual learning of the speaker's voice and speech pattern, this increases the overall learning benefit even in the noise conditions, perhaps by the high perceptual load mitigating the distractor effect of the noise. (2) Conversational speech heard with ARHL decreases the learning-practice benefit, with learning-practice benefits becoming much smaller relative to the clear speech style.

Implications

In summary, the results showed that when older adults listened to complex medical prescription instructions with "clear speech," (presented at audible levels through insertion earphones) their learning efficiency, immediate, and delayed memory performance improved relative to their performance when they listened with a normal conversational speech rate (presented at audible levels in sound field). This better learning and memory performance for clear speech listening was maintained even in the Noise group. When the speech was manipulated so that it was sufficiently discriminable in that it could be easily segregated into meaningful units (the clear speech technique), the presence of the irrelevant distractor - speech babble noise did not differentially affect memory performance. There was a weakly associated negative relationship between ARHL and delayed memory performance in this experiment. There were medium to large positive associations between delayed memory performance and working memory, executive control and lexical abilities; however, the magnitude of these effects were larger in the conversational listening compared to the clear listening condition. This finding indicates that explicit cognitive-linguistic abilities are correlated with delayed memory performance more so in sub-optimal or adverse listening conditions. It appears that those with strengths in cognitive-linguistic abilities are able to more efficiently compensate by re-allocating resources for discrimination and comprehension of the auditory-verbal message and still have sufficient resources for the secondary task of encoding the message in memory for later recall.

Further, these results suggest that the sources of interference (speaker, listener, and environment) may interact as follows. The auditory-verbal stimuli in the conversational speech relative to clear speech listening create a demand for more cognitive-linguistic resources to achieve successful decoding of the message. As a result, the listener's limited-capacity resources are re-allocated such that fewer resources are available for learning and encoding for later recall (effortfulness hypothesis). In addition, the finding that learning-practice effects were largest when clear speech was heard first, in both quiet (+3.25) and noise (+1.25), supports the hypothesis that a high perceptual load decreases the distractor effect, where a high perceptual load spoken with *conversational* style does not (Lavie, 2005). Perhaps then when older adults listen to conversational speech rate that is further degraded by ARHL (listener source of interference), the high perceptual load does not mitigate the distractor effect (environment issues - ambient noise/reverberation/babble), which then interferes with the on-line processing of the acoustic message. Results suggest that it is this environmental issue-the distraction (even milliseconds) from the online auditory temporal-spectral processing of the message that then requires those explicit cognitive-linguistic resources to decode the message, so that fewer resources are available for encoding for later recall.

Although the data showed a main effect of listening condition (conversational and clear) on learning and memory performance, the expectation was that the competition groups (Noise vs. Quiet) would be differentially affected by the listening condition resulting in an interaction of group with listening condition. This was not found, most likely because the noise was a between group variable and there were large variances in performance within the groups. However, an interaction of listening condition order with listening condition for the subgroups of 1st vs. 2nd listening conditions was evident reflecting a perceptual learning effect or adaptation of those who listened first in the clear speech.

In addition, the expectation was that the age-related auditory acuity deficit would be more strongly correlated with learning and memory performance for the two listening conditions. The expectation was that there would be a large negative effect of hearing-listening abilities, on learning and memory performance, with the magnitude of that effect being larger in the conversational compared to the clear listening condition (as a result of the signal). Perhaps the ARHL acuity deficit was completely corrected for by presenting the stimuli at the individual's MCL. If the presentation level was set at a fixed absolute hearing level (70 dB HL) this may have then resulted in the expected negative associations of greater ARHL and poorer delayed memory performance. It also could be because the groups' PTA4 reflected normal-to-moderate hearing loss at the higher frequencies. Use of MCL presentation level for a group of older adults with more severe, precipitously-sloping high-frequency hearing loss, would not have corrected for the hearing loss as completely. Perhaps then these ARHL factors would have negatively associated with delayed memory performance.

It is probable that once the stimuli were sufficiently audible, the level of temporal-spectral degrading did not reach a threshold or tipping point in which the added distortion from ARHL interacts with the processing of the message for successful recognition and comprehension. Instead it is the cognitive-linguistic abilities that are recruited as a compensatory process for successful recognition and encoding for later recall

(Bäckman and Dixon, 1992; Wild et al., 2012). The cognitive-linguistic scores significantly correlating with delayed memory performance with greater magnitudes in the conversational listening condition support this compensatory role of cognitive-linguistic abilities for adverse listening (Rudner et al., 2009).

Yet still the relative temporal-spectral manipulation of these two listening conditions might not have resulted in the conversational speech being sufficiently degraded. The temporal-spectral degrading of more typically produced conversational speech may not have been captured by this speaker's rendition. Since he was instructed to use articulation, rate and prosody for optimal clarity even for the original-conversation recording, and as a professionally trained singer and speaker, his normal conversational style is most likely comparable to citation-style speech. As Lam et al. (2012) demonstrated the instructions given to the speaker for the production of the passages affects the acoustic aspects and the intelligibility benefit (Krause and Braida, 2004, 2009; Lam et al., 2012). Citation–style speech production has been demonstrated to provide a larger intelligibility benefit than typically produced conversational speech and potentially only slightly less so from "clear speech technique" (Ferguson and Kewley-Port, 2007).

Nonetheless, enhancing the message by using a "clear speech" technique resulted in better learning and memory performance in two groups of older adults matched for age and ARHL. Additionally, the clear speech technique compared to conversational style speech reduced the negative impact that the competing noise had on learning and memory. Third, the finding that there was the largest learning effect on conversational speech as the second-listening condition after the clear speech listening condition was the first-listening condition of the experiment suggests greater perceptual learning or adaptation to the speaker's speech and voice pattern. This suggests that experience-dependent perceptual learning plays a role in facilitating or interfering with language processing and comprehension of a message and subsequent memory encoding.

Limitations and Future Directions

Ecologically valid methods and stimuli are preferred for understanding complex human behaviors in the context of real life, particularly for applicability and generalizability. However, there are inherent limitations such as fewer controls of latent variables, which may confound the results. For example, relevance, familiarity, and the subjective and objective importance of instructions can influence memory performance for older adults when processing larger quantities of information (Friedman et al., 2015). The vignettes in this study were developed to be intentionally relevant, important and generally familiar (medical-patch, puffer-inhaler). However, these variables were not actively manipulated in this study. Since relevance, importance and familiarity may interact with the listening conditions, future studies should consider manipulating and/or actively controlling for these variables. It is possible that these variables influence learning and memory more so in adverse listening conditions.

Another concern was the interaction between passage order, and listening condition order on immediate memory. It is possible that one passage may have lent itself to be spoken more "clearly" than another. In the future, experiments should use a more controlled method to spectrally and temporally enhance the stimuli such as a time-expansion technique (Tun, 1998; Peelle and Wingfield, 2005). Also, to examine whether a more substantial manipulation of the temporal-spectral aspect of the stimuli interacts with ARHL, either more typically spoken conversational speech or a time-compressed technique could be employed. Additionally, using the competition as a within subject variable instead of as a between subject variable will capture the degree to which the ARHL interacts with the noise and further increases listening effort for language processing and comprehension of the message. Finally, by using a more controlled enhancement such as expanded speech in quiet this manipulation would more closely resemble the experience that the younger adult has when listening. Then younger and older participant group's learning and memory performance could be compared in the two listening conditions (time-compressed with noise and time-expanded in quiet). Those aspects that mimic ARHL should then result in poorer learning and memory performance, and those that mimic younger listening should result in better learning and memory performance for both groups. With a within-subject research design one can then examine the relationships of hearing-listening factors and cognitive-linguistic characteristics on the learning and memory performance during the two listening conditions.

Acknowledgments

This work was funded by the Newfoundland and Labrador Healthy Aging Research Program, Doctoral Dissertation Award (2012), Doctoral Research Grant (2011); and the Canadian Institutes of Health Research Gold award (2011).

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00778/abstract

References

American National Standards Institute. (2004). *Specification for Audiometers. ANSI S3.6-2004, ANSI.* Melville, NY: Acoustical Society of America (ASA).

Amichetti, N. M., Stanley, R. S., White, A. G., and Wingfield, A. (2013). Monitoring the capacity of working memory: executive control and effects of listening effort. *Mem. Cognit.* 41, 839–849. doi: 10.3758/s13421-013-0302-0

Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). Aging affects neural precision of speech encoding. *J. Neurosci.* 32, 14156–14164. doi: 10.1523/JNEUROSCI.2176-12.2012

Anderson, S., Parbery-Clark, A., Yi, H., and Kraus, N. (2011). A neural basis of speech-in-noise perception in older adults. *Ear Hear.* 32, 750–757. doi: 10.1097/AUD.0b013e31822229d3

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hear. Res.* 300, 18–32. doi: 10.1016/j.heares.2013.03.006

Bäckman, L., and Dixon, R. A. (1992). Psychological compensation: a theoretical framework. *Psychol. Bull.* 112, 259–283. doi: 10.1037/0033-2909.112.2.259

Baker, R. E., and Bradlow, A. R. (2009). Variability in word duration as a function of probability, speech style, and prosody. *Lang. Speech* 52, 391–413. doi: 10.1177/0023830909336575

Baldwin, C. L., and Ash, I. K. (2011). Impact of sensory acuity on auditory working memory span in young and older adults. *Psychol. Aging* 26, 85–91. doi: 10.1037/a0020360

Baltes, P. B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging. *Psychol. Aging* 12, 12–21. doi: 10.1037/0882-7974.12.1.12

Boersma, P., and Weenink, D. (2014). *Praat: Doing Phonetics by Computer Version 5.3.63*. University of Amsterdam, Amsterdam. Available online at: http://www.praat.org/

Bradlow, A., Kraus, N., and Hayes, E. (2003). Speaking clearly for children with learning disabilities: sentence perception in noise. *J. Speech Lang. Hear. Res.* 46, 80–97. doi: 10.1044/1092-4388(2003/007)

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol. Assess.* 6, 284–290. doi: 10.1037/1040-3590.6.4.284

Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 115–159. doi: 10.1037/0033-2909.112.1.155

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: a methodological review and user's guide. *Psychon. Bull. Rev.* 12, 769–786. doi: 10.3758/BF03196772

Craik, F. I. M., Anderson, N. D., Kerr, S. S., and Li, K. Z. H. (1995). "Memory changes in normal ageing," in *Handbook of Memory Disorders,* eds A. D. Baddeley, B. A. Wilson, and F. N. Watts (New York, NY: Wiley), 211–241.

Craik, F. I. M., and Bialystok, E. (2006). Planning and task management in older adults: cooking breakfast. *Mem. Cognit.* 34, 1236–1249. doi: 10.3758/BF03193268

Crum, R. M., Anthony, J. C., Bassett, S. S., and Folstein, M. F. (1993). Population-based norms for the mini-mental state examination by age and educational level. *J. Am. Med. Assoc.* 269, 2386–2391. doi: 10.1001/jama.1993.03500180078038

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learning Verbal Behav.* 19, 450. doi: 10.1016/S0022-5371(80)90312-6

DiDonato, R. (2014). *Effortful and Effortless Listening: How Age-Related Hearing Loss and Cognitive Abilities Interact and Influence Memory Performance in Older Adults*. Ph.D. thesis, Memorial University of Newfoundland.

Fant, G. (1973). *Speech Sounds and Features*. Cambridge, MA: The MIT Press.

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Ferguson, S. H. (2012). Talker differences in clear and conversational speech: vowel intelligibility for older adults with hearing loss. *J. Speech Lang. Hear. Res.* 55, 779–790. doi: 10.1044/1092-4388(2011/10-0342)

Ferguson, S. H., and Kewley-Port, D. (2007). Talker differences in clear and conversational speech: acoustic characteristics of vowels. *J. Speech Lang. Hear. Res.* 50, 1241–1255. doi: 10.1044/1092-4388(2007/087)

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Friedman, M. C., McGilliray, S., Murayama, K., and Castel, A. D. (2015). Memory for medication side effects in younger and older adults: the role of subjective and objective importance. *Mem. Cognit.* 43, 206–215 doi: 10.3758/s13421-014-0476-0

Gilbert, R. C., Chandrasekaran, B., and Smiljanic, R. (2014). Recognition memory in noise for speech of varying intelligibility. *J. Acoust. Soc. Am.* 135, 389–399. doi: 10.1121/1.4838975

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous Speech*. New York, NY: Academic Press.

Goodglass, H., Kaplan, E., and Barresi, B. (2001). *The Assessment of Aphasia and Related Disorders, 4th Edn.* Baltimore, MD: Lippincott Williams and Wilkins.

Gordon-Salant, S., and Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *J. Speech Lang. Hear. Res.* 36, 1276–1285. doi: 10.1044/jshr.3606.1276

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor. Quant. Methods Psychol.* 8, 23–34. Available online at: http://www.tqmp.org/RegularArticles/vol08-1/p023/p023.pdf

Hasher, L., and Zacks, R. T. (1979). Automatic and effortful processes in memory. *J. Exp. Psychol.* 108, 356–388. doi: 10.1037/0096-3445.108.3.356

Hasher, L., and Zacks, R. T. (1988). "Working memory, comprehension, and aging: a review and a new view," in *The Psychology of Learning and Motivation: Advances in Research and Theory,* Vol. 22, ed G. H. Bower (San Diego, CA: Academic Press), 193–225.

Hasher, L., Zacks, R. T., and May, C. P. (1999). "Inhibitory control, circadian arousal, and age," in *Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application,* eds D. Gopher and A. Koriat (Cambridge, MA: The MIT Press), 653–675. Available online at: http://www.psych.utoronto.ca/users/hasherlab/abstracts/hasher_zacks_may99.htm

Hausknect, J. P., Halpert, J. A., DiPaolo, N. T., and Moriarty, G. M. O. (2006). Retesting in selection: a meta-analysis of coaching and practice effects for tests of cognitive ability. *J. Appl. Psychol.* 92, 373–385. doi: 10.1037/0021-9010.92.2.373

Heinrich, A., Schneider, B. A., and Craik, F. I. M. (2008). Investigating the influence of continuous babble on auditory short-term memory performance. *Q. J. Exp. Psychol.* 61, 735–751. doi: 10.1080/17470210701402372

Helfer, L. S., and Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *J. Am. Acad. Audiol.* 20, 264–271. doi: 10.3766/jaaa.20.4.6

Humes, L. E. (2008). Aging and speech communication: peripheral, central-auditory, and cognitive factors affecting the speech-understanding problems of older adults. *ASHA Lead.* 13, 10–33. doi: 10.1044/leader.FTR1.13052008.10

Humes, L. E., Dubno, J. R., Gordon-Salant, S., Lister, J. J., Cacace, A. T., Cruickshanks, K. J., et al. (2012). Central presbycusis: a review and evaluation of the evidence. *J. Am. Acad. Audiol.* 23, 635–666. doi: 10.3766/jaaa.23.8.5

John, A. B., Hall, J. W., and Kreisman, B. M. (2012). Effects of advancing age and hearing loss on gaps-in-noise test performance. *Am. J. Audiol.* 21, 242–250. doi: 10.1044/1059-0889(2012/11-0023)

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual difference in working memory. *Psychol. Rev.* 99:122. doi: 10.1037/0033-295X.99.1.122

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kaplan, E., Goodglass, H., and Weintraub, S. (2001). *Boston Naming Test, 2nd Edn.* Austin, TX: Pro-ed.

Katz, J. (ed.). (1978). *Handbook of Clinical Audiology, 2nd Edn*. Baltimore, MD: Williams and Wilkins.

Keisler, A., and Willingham, D. T. (2007). Non-declarative sequence learning does not show savings in relearning. *Hum. Mov. Sci.* 26, 247–256. doi: 10.1016/j.humov.2007.01.003

Killion, M. (2002). New thinking on hearing in noise: a generalized articulation index. *Semin. Hear.* 23, 57. doi: 10.1055/s-2002-24976

Killion, M., Niquette, P., Gudmundsen, G. I., Revit, L., and Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 116, 2395–2405. doi: 10.1121/1.1784440

Konkle, D. F., Beasley, D. S., and Bess, F. H. (1977). Intelligibility of time-altered speech in relation to chronological aging. *J. Speech Lang. Hear. Res.* 20, 108–115. doi: 10.1044/jshr.2001.108

Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605. doi: 10.1038/nrn2882

Krause, J., and Braida, L. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *J. Acoust. Soc. Am.* 115, 362–378. doi: 10.1121/1.1635842

Krause, J., and Braida, L. (2009). Evaluating the role of spectral and envelop characteristics in the intelligibility advantage of clear speech. *J. Acoust. Soc. Am.* 125, 3346–3357. doi: 10.1121/1.3097491

Lam, J., Tjaden, K., and Wilding, G. (2012). Acoustics of clear speech: effect of instruction. *J. Speech Lang. Hear. Res.* 55, 1807–1821. doi: 10.1044/1092-4388(2012/11-0154)

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends Cogn. Sci.* 9, 75. doi: 10.1016/j.tics.2004.12.004

Lavie, N., and DeFockert, J. W. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Percept. Psychophys.* 65, 202. doi: 10.3758/BF03194795

Light, L. L. (1991). Memory and aging: four hypotheses in search of data. *Annu. Rev. Psychol.* 42, 333. doi: 10.1146/annurev.ps.42.020191.002001

Mattys, S. L., Brooks, J., and Cooke, M. (2009). Recognizing speech under a processing load: dissociating energetic from informational factors. *Cogn. Psychol.* 59, 203–243. doi: 10.1016/j.cogpsych.2009.04.001

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006

Mattys, S. L., and Scharenborg, O. (2014). Phoneme categorization and discrimination in younger and older adults: a comparative analysis of perceptual, lexical and attentional factors. *Psychol. Aging* 29, 150–162. doi: 10.1037/a0035387

Mattys, S. L., and Wiget, L. (2011). Effects of cognitive load on speech recognition. *J. Mem. Lang.* 65, 145–160. doi: 10.1016/j.jml.2011.04.004

McCoy, S., Tun, P., Cox, L., Colangelo, M., Stewart, R., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol. A* 58, 22–33. doi: 10.1080/02724980443000151

Mueller, J. A., and Dollaghan, C. (2013). A systematic review of assessments for identifying executive function impairments in adults with acquired brain injury. *J. Speech Lang. Hear. Res.* 56, 1051–1064. doi: 10.1044/1092-4388(2012/12-0147)

Newby, H. (ed.). (1979). *Audiology, 4th Edn*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Newman, C. W., Weinstein, B. E., Jacobson, G., and Hug, G. (1991). Test-retest reliability of the hearing handicap inventory for adults. *Ear Hear.* 12, 355–357. doi: 10.1097/00003446-199110000-00009

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.* 52, 433–441. doi: 10.3109/14992027.2013.776181

Parbery-Clark, A., Anderson, S., Hittner, E., and Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiol. Aging* 33, 1483–1481. doi: 10.1016/j.neurobiolaging.2011.12.015

Parbery-Clark, A., Skoe, E., Lam, C., and Kraus, N. (2009). Musician enhancement for speech in noise. *Ear Hear.* 30, 653–661. doi: 10.1097/AUD.0b013e3181b412e9

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., et al. (1996). Mediators of long-term memory performance across the life span. *Psychol. Aging* 11, 621–637. doi: 10.1037/0882-7974.11.4.621

Peelle, J., and Wingfield, A. (2005). Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech. *J. Exp. Psychol.* 31:1315. doi: 10.1037/0096-1523.31.6.1315

Picheny, M. A., Durlach, N. I., and Braida, L. D. (1985). Speaking clearly for the hard of hearing I. Intelligibility differences between clear and conversational speech. *J. Speech Lang. Hear. Res.* 28, 96–103. doi: 10.1044/jshr.2801.96

Pichora-Fuller, K., Schneider, B., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593. doi: 10.1121/1.412282

Rabbitt, P. (1968). Channel-capacity, intelligibility and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158

Rabbitt, P. (1990). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol. Suppl.* 111, 167–175. doi: 10.3109/00016489109127274

Ritter, F. E., Reifers, A., Klein, L. C., Quigley, K., and Schoelles, M. (2004). "Using cognitive modeling to study behavior moderators: pre-task appraisal and anxiety," in *Human Factors and Ergonomics Society Annual Meeting Proceedings* (Santa Monica, CA: Human Factors and Ergonomics Society), 48, 2121–2125. doi: 10.1177/154193120404801709

Ritter, F. E., and Schooler, L. J. (2001). "The learning curve," in *International Encyclopedia of the Social and Behavioral Sciences,* eds W. Kintch, N. Smelser, and P. Baltes (Amsterdam: Pergamon), 8602–8605. doi: 10.1016/B0-08-043076-7/01480-7

Roach, A., Schwartz, M. R., Martin, N., Grewal, R. S., and Brecher, A. (1996). The Philadelphia naming test: scoring and rationale. *Clin. Aphasiol.* 24, 121–133.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol. 47*, S99-S105. doi: 10.1080/14992020802301167

Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 49, 263–269. doi: 10.4103/1463-1741.70505

Rosen, S. (1992). "Temporal information in speech: acoustic, auditory and linguistic aspects," in *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences,* Vol. 336, eds R. P. Carlyon, C. J. Darwin, and I. J. Russell (New York, NY: Clarendon Press/Oxford University Press), 367–373.

Rudner, M., Foo, C., Rönnberg, J., and Lunner, T. (2009). Cognition and aided speech recognition in noise: specific role for cognitive factors following nine-weeks experience with adjusted compression settings in hearing aids. *Scand. J. Psychol.* 50, 405–418. doi: 10.1111/j.1467-9450.2009.00745.x

Salthouse, T., A (2010). *Major Issues in Cognitive Aging*. Oxford; New York, NY: Oxford University Press.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403

Schaefer, S. (2014). The ecological approach to cognitive-motor dual-tasking: findings on the effects of expertise and age. *Front. Psychol.* 5:1167. doi: 10.3389/fpsyg.2014.01167

Schneider, B., and Pichora-Fuller, K. (2000). "Implications of perceptual deterioration for cognitive aging research," in *The Handbook of Aging and Cognition, 2nd Edn.* eds F. I. M. Craik, and T. Salthouse (Mahwah, NJ: Lawrence Erlbaum Associates, Inc.), 155.

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428. doi: 10.1037/0033-2909.86.2.420

St Clair-Thompson, H. L. (2010). Backwards digit recall: a measure of short-term memory or working memory? *Eur. J. Cogn. Psychol.* 22, 286–296. doi: 10.1080/09541440902771299

St Clair-Thompson, H., and Sykes, S. (2010). Scoring methods and the predictive ability of working memory tasks. *Behav. Res. Methods* 42, 969–975. doi: 10.3758/BRM.42.4.969

Stewart, R., and Wingfield, A. (2009). Hearing loss and cognitive effort in older adults' report accuracy for verbal materials. *J. Am. Acad. Audiol.* 20, 147–154. doi: 10.3766/jaaa.20.2.7

Surprenant, A. (1999). The effect of noise on memory for spoken syllables. *Int. J. Psychol.* 34, 328–333. doi: 10.1080/002075999399648

Surprenant, A. (2007). Effects of noise on identification and serial recall of nonsense syllables in older and younger adults. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 14, 126–143. doi: 10.1080/13825580701217710

Tun, P. A. (1998). Fast noisy speech: age differences in processing rapid speech with background noise. *Psychol. Aging* 13, 424–434. doi: 10.1037/0882-7974.13.3.424

Tun, P. A., O'Kane, G., and Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychol. Aging* 17, 453–467. doi: 10.1037/0882-7974.17.3.453

Tun, P., McCoy, S., and Wingfield, A. (2009). Aging, hearing acuity, and the attentional cost of effortful listening. *Psychol. Aging* 24, 761. doi: 10.1037/a0014802

Ventry, I., and Weinstein, B. E. (1982). The hearing handicap inventory for the elderly: a new tool. *Ear Hear.* 3, 128–134. doi: 10.1097/00003446-198205000-00006

Verhaeghen, P., and Salthouse, T. A. (1997). Meta-analyses of age–cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. *Psychol. Bull.* 122, 231–249. doi: 10.1037/0033-2909.122.3.231

Wambacq, I. J. A., Koehnke, J., Besing, J., Romei, L. L., DePierro, A., and Cooper, D. (2009). Processing interaural cues in sound segregation by young and middle-aged brains. *J. Am. Acad. Audiol.* 20, 453–458. doi: 10.3766/jaaa.20.7.6

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale Revised*. San Antonio, TX: The Psychological Corporation.

Wild, C., Yusuf, A., Wilson, D., Peelle, J., Davis, M., and Johnsrude, I. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *J. Neurosci.* 32, 14010–14021. doi: 10.1523/JNEUROSCI.1528-12.2012

Wingfield, A., and Ducharme, J. L. (1999). Effects of age and passage difficulty on listening-rate preferences for time-altered speech. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 54B, P199–P202. doi: 10.1093/geronb/54B.3.P199

Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *J. Am. Acad. Audiol.* 17, 487–497. doi: 10.3766/jaaa.17.7.4

Wingfield, A., Poon, L. W., Lombardi, L., and Lowe, D. (1985). Speed of processing in normal aging: effects of speech rate, linguistic structure, and processing time. *J. Gerontol.* 40, 579–585. doi: 10.1093/geronj/40.5.579

Wingfield, A., Tun, P. A., Koh, C. K., and Rosen, M. J. (1999). Regaining lost time: adult aging and the effect of time restoration on recall of time-compressed speech. *Psychol. Aging* 14, 380–389. doi: 10.1037/0882-7974.14.3.380

Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Working Group on Speech Understanding and Aging and the Committee on Hearing, Bioacoustics and Biomechanics (CHABA). (1988). Speech understanding and aging. *J. Acoust. Soc. Am.* 83, 859–895. doi: 10.1121/1.395965

World Health Organization Prevention of Blindness and Deafness (PBD) Program. (2014). *Prevention of Deafness and Hearing Impaired Grades of Hearing Impairment*. Available online at: http://www.who.int/pbd/deafness/hearing_impairment_grades/en/index.html (Accessed January 06, 2014).

Zacks, R. T., and Hasher, L. (1994). "Directed ignoring: inhibitory regulation of working memory," in *Inhibitory Processes in Attention, Memory, and Language,* eds D. Dagenbach and T. H. Carr (San Diego, CA: Academic Press), 241–264.

Zacks, R. T., and Hasher, L. (1997). Cognitive gerontology and attentional inhibition: a reply to Burke and McDowd. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 52, 274–284. doi: 10.1093/geronb/52B.6.P274

Zacks, R. T., Hasher, L., and Li, K. Z. H. (2000). "Human memory," in *The Handbook of Aging and Cognition, 2nd Edn.,*eds F. I. M. Craik, and T. A. Salthouse (Mahwah, NJ: Lawrence Erlbaum Associates), 293–357.

Zendel, B. R., and Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* 27, 410–417. doi: 10.1037/a0024816

Zendel, B. R., and Alain, C. (2013). The influence of lifelong musicianship on neurophysiological measures of concurrent sound segregation. *J. Cogn. Neurosci.* 25, 503–515. doi: 10.1162/jocn_a_00329

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 DiDonato and Surprenant. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 19 February 2015
doi: 10.3389/fpsyg.2015.00177

Hearing loss impacts neural alpha oscillations under adverse listening conditions

Eline B. Petersen[1,2,3]*, Malte Wöstmann[4,5], Jonas Obleser[5], Stefan Stenfelt[2,3] and Thomas Lunner[1,3]

[1] Eriksholm Research Centre, Snekkersten, Denmark
[2] Technical Audiology, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden
[3] Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden
[4] International Max Planck Research School on Neuroscience of Communication, Leipzig, Germany
[5] Max Planck Research Group "Auditory Cognition," Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Edited by:
Claude Alain, Rotman Research Institute, Canada

Reviewed by:
Jochen Kaiser, Johann Wolfgang Goethe University, Germany
Yi Du, Rotman Research Institute – Baycrest Centre for Geriatric Care, Canada

****Correspondence:***
Eline B. Petersen, Eriksholm Research Centre, Rørtangvej 20, 3070 Snekkersten, Denmark
e-mail: ebp@eriksholm.com

Degradations in external, acoustic stimulation have long been suspected to increase the load on working memory (WM). One neural signature of WM load is enhanced power of alpha oscillations (6–12 Hz). However, it is unknown to what extent common internal, auditory degradation, that is, hearing impairment, affects the neural mechanisms of WM when audibility has been ensured via amplification. Using an adapted auditory Sternberg paradigm, we varied the orthogonal factors memory load and background noise level, while the electroencephalogram was recorded. In each trial, participants were presented with 2, 4, or 6 spoken digits embedded in one of three different levels of background noise. After a stimulus-free delay interval, participants indicated whether a probe digit had appeared in the sequence of digits. Participants were healthy older adults (62–86 years), with normal to moderately impaired hearing. Importantly, the background noise levels were individually adjusted and participants were wearing hearing aids to equalize audibility across participants. Irrespective of hearing loss (HL), behavioral performance improved with lower memory load and also with lower levels of background noise. Interestingly, the alpha power in the stimulus-free delay interval was dependent on the interplay between task demands (memory load and noise level) and HL; while alpha power increased with HL during low and intermediate levels of memory load and background noise, it dropped for participants with the relatively most severe HL under the highest memory load and background noise level. These findings suggest that adaptive neural mechanisms for coping with adverse listening conditions break down for higher degrees of HL, even when adequate hearing aid amplification is in place.

Keywords: alpha oscillations, hearing loss, hearing aid, cognition, working memory

INTRODUCTION

Adverse listening conditions are common in everyday life. Auditory distractions and signal degradations increase demands on attention and working memory (WM; Shinn-Cunningham and Best, 2008). WM describes the system for temporary storage and processing of information to perform a cognitive task (Baddeley, 1986). Any degradation of the sensory auditory input requires increased WM involvement to successfully interpret the stimuli (Rönnberg et al., 2008; Stenfelt and Rönnberg, 2009). Auditory stimuli can be degraded by external factors, often occurring in the form of background noise, in which case WM is engaged to extract useful information from the auditory input (Pichora-Fuller, 2003). However, auditory processing can also be disrupted by internal degradation, such as sensorineural hearing loss (HL). To alleviate this internal degradation of the auditory input, people suffering from HL are typically treated with hearing aids. The purpose of a hearing aid is to amplify the auditory input to make sounds audible and consequently reduce the internal auditory degradation, which theoretically should release WM resources (sometimes referred to as lowered cognitive load; Lunner, 2003). Here, we tested whether HL affects brain signatures of WM involvement in an adverse listening paradigm.

The power of neural oscillations in the alpha frequency band (liberally defined as 6–12 Hz) has been found to increase with WM load (Jensen et al., 2002). According to the functional inhibition framework (Klimesch et al., 2007; Jensen and Mazaheri, 2010), alpha oscillations indicate the inhibition of currently task-irrelevant brain regions and/or cognitive processes to prevent interference with task-relevant cognitive processing (Bonnefond and Jensen, 2012). Although alpha power modulations have been found for external degradation of auditory signals (van Dijk et al., 2010; Obleser and Weisz, 2012; Obleser et al., 2012; Becker et al., 2013; Scharinger et al., 2014; Wöstmann et al., 2015), it is currently unknown how the internal degradation of auditory input through HL affects neural alpha dynamics (Strauß et al., 2014). There is good evidence from behavioral studies that HL negatively affects cognitive operations on the speech signal (McCoy et al., 2005; Wingfield et al., 2005, 2006). These findings support the hypothesis put forward by Rabbitt (1991), stating that adverse listening conditions require the allocation of more cognitive resources, which

could otherwise be used for more task-relevant cognitive processing, such as storing information. Thus, external (acoustic), and internal (auditory) degradations are assumed to trigger a higher degree of WM involvement during the encoding of task-relevant stimuli, leaving fewer cognitive resources for the storage, and processing of information in the WM (Lunner et al., 2009; Van Engen and Peelle, 2014). Here, we tested whether HL impacts behavioral performance and neural mechanisms even when it is treated with individually fitted hearing aids.

A well-established experimental paradigm to test WM demands is the Sternberg paradigm (Sternberg, 1966). Participant's task is to encode and retain a number of items to compare them to a subsequent probe. Although the Sternberg paradigm was originally developed as a visual WM task, it has since been adapted to test auditory WM (e.g., Rojas et al., 2000; Leiberg et al., 2006). The test incorporates a short stimulus-free delay period between the encoding and the probe presentation, during which the participants are to retain the presented stimuli in memory. This stimulus-free delay period is of special interest in neuroimaging studies, because neural responses measured in this time period are thought to reflect WM processes independent of the sensory stimulation itself. During stimuli presentation, the processes of auditory encoding and memory storage are not easily separated, contrary to the delay period where there is no sensory input and the only task is to retain the stimuli in memory and restore inadequately encoded items. A number of studies have found that increased memory load (i.e., increasing the number of items to be remembered) was associated with enhanced alpha power over central and parietal recording sites during the delay period (Jensen et al., 2002; Leiberg et al., 2006; Obleser et al., 2012). Critically, Obleser et al. (2012) recently found that alpha power in the delay period was not only enhanced with an increasing number of to-be-remembered items, but with the acoustic degradation of the items.

In the present study, a version of the Sternberg test modified by Obleser et al. (2012) was applied to investigate the effects of varying memory load and the level of background noise on alpha oscillations measured by electroencephalogram (EEG) recording. We tested older participants with varying degrees of HL. In line with prior studies, we expected decreased task performance with higher memory load and higher levels of background noise. We hypothesized that alpha power would increase with the severity of HL, suggesting that internal auditory degradations increase the load on neural WM mechanisms in speech processing. Furthermore, it was of interest whether such increased expenditure of cognitive resources would reach a limit and break down (i.e., reminiscent of the CRUNCH hypothesis put forward by Reuter-Lorenz and Cappell, 2008) in listeners with the most severe HL and/or under highest task demands (i.e., highest memory load and most severe background noise).

MATERIALS AND METHODS

PARTICIPANTS

Twenty-nine native Swedish speaking participants (16 females, age range: 62–86 years, mean age 72.2 years), recruited from the audiology clinic at the University Hospital of Linköping in Sweden, participated in this study. Participants were recruited to show large inter-individual variability of auditory pure-tone thresholds. Participants were grouped according to their pure-tone average (PTA), across 0.5, 1, 2, 4, and 8 kHz into three groups of HL (no/mild/moderate HL). The hearing threshold at 8 kHz was included in the PTA since sensitivity loss at higher frequencies is known to accompany age-related HL (CHABA, 1988). Separate one-way ANOVAs showed no difference in age between groups ($p = 0.114$), but a significant difference in HL ($p < 0.001$), with Fisher's Least Significant Difference (LSD) *post hoc* analysis showing significant differences between the three groups (all $p < 0.001$). Participant information is shown in **Table 1** and **Figures 1C,D**.

Participants all gave informed consent and were given no financial compensation for their participation. The study was approved by the regional ethical review in Linköping, Sweden and conformed with the Helsinki Declaration of Ethical Principles for Medical Research Involving Human Subjects.

EXPERIMENTAL DESIGN

Speech materials

The stimuli consisted of the monosyllabic Swedish digits "0," "1," "2," "3," "5," "6," and "7," spoken by a female talker and recoded in a soundproof booth at a sampling rate of 22.05 kHz. For a natural co-articulation, the digits were recorded as triplets. The triplets were adjusted to the same root-mean-square (RMS) level, and then the first digit was extracted without silent intervals before and after each waveform, resulting in an average digit duration of 677 ms (SD: 103 ms). The recordings were originally used for the Swedish digit triplets test (Drullman et al., 2005; Larsby et al., 2011).

The final audio files were generated by adding speech-shaped noise to the digits at the individualized SNR levels (see below). Due to the short duration of the spoken digits acceptable speech-shaped noise could not be generated based on the spectrum of the digits. The speech-shaped noise was taken from the Dantale II test, a standardized speech intelligibility test (Wagener et al., 2003). Speech-shaped noise is random stationary broadband noise, with the same long-term average frequency spectrum as natural speech.

Stimulus presentation

All participants were wearing Agil hearing aids (Oticon A/S, Smørum, Denmark) with individual quasi-linear amplification. The quasi-linear amplification accounts for the audibility of soft (inaudible speech) sounds by incorporating a fast-acting gain adjustment at the onset of the presented sounds and maintaining this gain throughout the presentation of the sounds with a very slow-acting gain adjustment (for details see Simonsen and Behrens, 2009). No changes were made to the time constant throughout the sound presentation, and the hearing aid amplification can be considered linear, meaning that the hearing aid output intensity increased at the same rate as the intensity of the acoustic input. The noise reduction algorithm and volume control normally available on these hearing aids were disabled during the entire experimental session.

All auditory stimuli were presented directly through the hearing aids using the Direct Audio Input (DAI). The experiment was conducted in an electrically shielded soundproof booth. Visual

Table 1 | Participant information.

	Hearing threshold range [dB HL]	Pure-tone average (PTA) [dB HL]	Age [years]	No. of females
No hearing loss (HL; $n = 8$)	0–25	22.3 (7.1)	68.8 (4.6)	4
Mild HL ($n = 11$)	25–50	42.1 (8.4)	72.5 (5.5)	7
Moderate HL ($n = 10$)	50–80	63.7 (5.2)	74.6 (6.4)	5
Total ($n = 29$)		44.1 (17.9)	72.2 (5.8)	16

Participants grouped according to their HL (no/mild/moderate HL; first column), defined based on three ranges of hearing thresholds (second column). Values in parentheses indicate one standard deviation. Average hearing threshold levels for the three groups across 0.5, 1, 2, 4, and 8 KHz are shown in the third column. Columns four and five list participants' mean age and number of females in the three groups, respectively. The bottom row shows average data across the entire sample of participants.

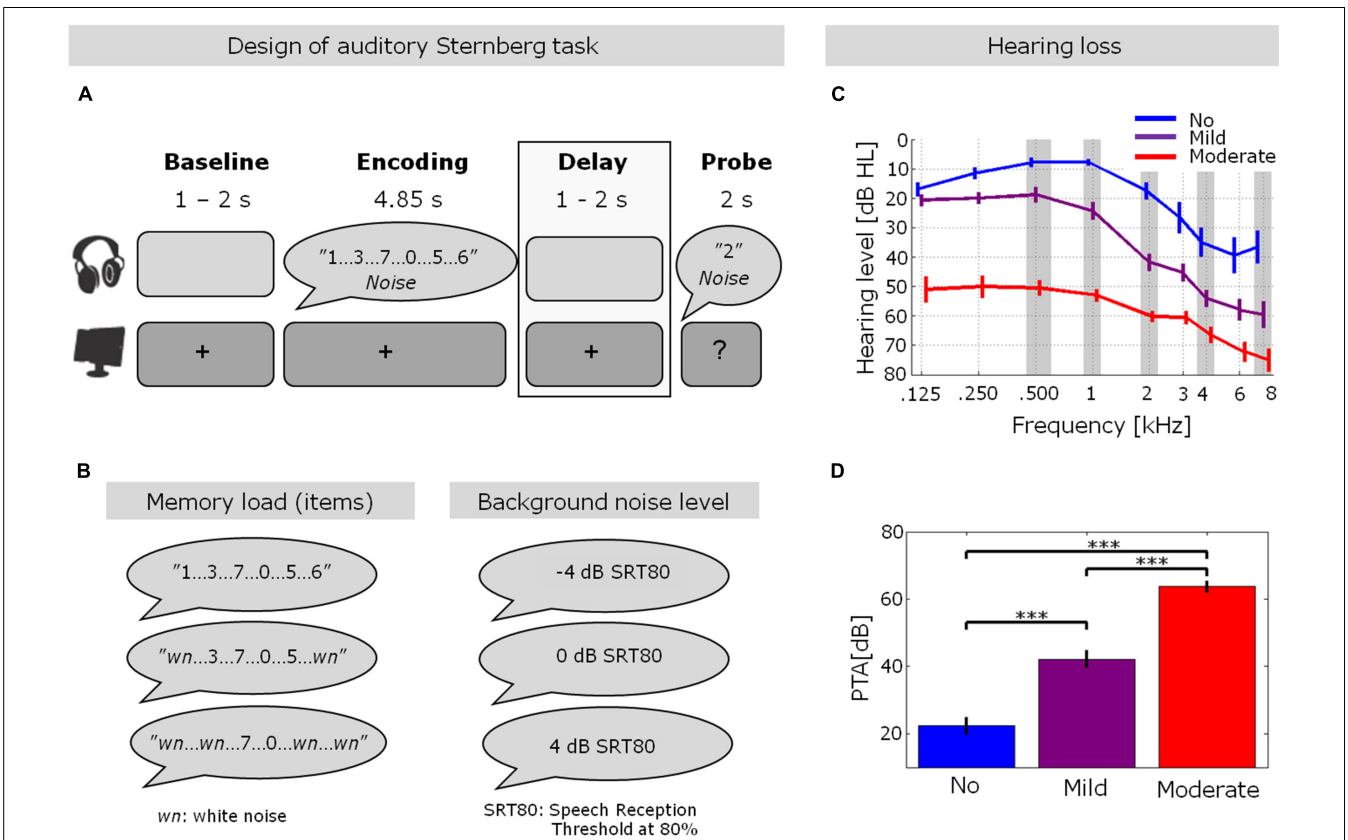


FIGURE 1 | Hearing thresholds and experimental design. (A,B) Trial design in the auditory Sternberg task. After an initial silent baseline period, participants were presented with a varying number of spoken digits (2, 4, or 6; see *Experimental procedure* for details) embedded in three different individually adjusted background noise levels (–4, 0, or 4 dB relative to the individual speech reception threshold at 80%, SRT80). After a silent delay period, participants indicated whether a probe digit was presented during the encoding. The gray box highlights the stimulus-free delay period, which was the focus of the EEG data analysis in the present study. **(C)** Pure-tone hearing thresholds for the three hearing loss (HL) groups (blue: no HL, purple: mild HL, red: moderate HL). Error bars indicate ±1 SEM. **(D)** Pure-tone average across frequencies highlighted with gray shading in **(C)** (0.5, 1, 2, 4, and 8 kHz) for the three groups of HL (***$p < 0.001$; one-way ANOVA with Fisher's LSD *post hoc* analyses). These PTA values are also shown in **Table 1**. Error bars indicate ±1 SEM. The figure is adapted from Obleser et al. (2012).

cues and instructions were presented on a 1280 by 1024 resolution screen, with the participants positioned 1 m from the screen.

Individual adjustments of SNR levels

To ensure equal intelligibility of the stimulus materials for all participants despite large inter-individual differences in hearing thresholds (see **Figures 1C,D**; **Table 1**), the background noise levels were individually adjusted. To this end, participants listened to and repeated 40 spoken sentences from the Swedish version of hearing in noise test (HINT; Hällgren et al., 2006). The output presentation level was 70 dB SPL, which was presented through the DAI of the hearing aids and amplified according to the individual

audiograms. In an adaptive tracking procedure (Levitt, 1971), we determined the background noise level (measured as the signal to noise ratio between speech and background noise) at which each participant was able to repeat 80% of the words in a sentence. This value for an individual participant will be referred to as the Speech Reception Threshold (SRT) of 80% (denoted 0 dB SRT80). In the Sternberg test, the individual 0 dB SRT80 level was used as the intermediate background noise level for the participant in question. The lower and higher background noise levels were generated by raising or lowering the SNR by 4 dB from the obtained 0 dB SRT80, denoted 4 dB SRT80 and –4 dB SRT80, respectively. To maintain a constant overall intensity level of the stimuli played from the presentation computer at ∼70 dB SPL, both the level of the signal (i.e., the digits) and the level of the background noise were adjusted. For instance, for the 4 dB SRT80 condition, the noise level was lowered by 2 dB in intensity, and the signal level was raised by 2 dB relative to the 0 dB SRT80.

Experimental procedure

After the individual adjustment of SNR levels, the actual experiment was performed. An auditory version of the Sternberg paradigm (Sternberg, 1966), inspired by Obleser et al. (2012), was used, employing a 3 × 3 design of the orthogonal factors memory load (2, 4, or 6 digits to be remembered) and background noise level (4 dB, 0 dB, or –4 dB relative to the individual level at which 80% of the words were correctly recalled in noise). Each trial started with the presentation of a central fixation cross for 1–2 s (randomly varied duration), followed by the encoding phase, in which 2, 4, or 6 digits were presented in speech-shaped noise (**Figures 1A,B**). The noise onset always preceded the onset of the first digit by 50 ms to avoid masking of the first digit by the noise onset. In trials with two and four digits, flanking sounds of white noise, at the same intensity level as the spoken digits, were presented to always ensure the presentation of six sounds. The sounds (digits and flanking noises) were presented with an onset-to-onset stimulus interval of 0.8 s, resulting in a total encoding time of 4.85 s, after which the noise was also terminated.

The encoding was followed by a stimulus-free delay period, in which the participants were to retain the presented digits in their memory. The delay phase had a duration of 1–2 s (randomly varied). Lastly, a probe digit was presented in the same background noise level as during the encoding interval. Again, the noise started 50 ms prior to the probe digit. During this 50 ms interval, the fixation cross changed to a question mark, signaling that the participants were to indicate, via a button press on a response box, whether the probe digit appeared in the encoding phase (response window of 2 s). Participants were not instructed to use any particular finger(s) for pressing the response buttons, nor were the button positions varied between participants. If participants required more than 2 s to respond, they were instructed to be faster on the next trial and informed that no response was recorded. Feedback was given after each trial, consisting of either 'correct,' 'incorrect,' or 'no answer registered, please answer faster.' In half of the trials, the probe digit appeared during encoding.

Trials for the nine conditions in the 3 (memory load) × 3 (background noise level) design were presented in 10 blocks. Due to the length of the test, the 10 test blocks were separated into two recordings of five blocks. Each recording lasted ∼45 min with a break of 15 min between the two recordings. Each recording was initiated with a training block of 11–25 trials from all nine conditions. Each test block consisted of a minimum of 18 trials with 2 trials for each condition, presented in a randomized order. The actual number of trials per block was determined by the number of unanswered trials. That is, for each trial in which no answer was registered due to a response time longer than 2 s, an extra trial was added to the block. Overall, 20 trials with registered answers were recorded in each condition for each participant.

EEG RECORDING AND PREPROCESSING

The EEG was recorded using an EGI system (Electrical Geodesic Inc., Eugene, OR, USA) with 128 Ag/Ag-Cl channels. Six occipital and one central electrode were disconnected from the electrode net and used for other physiological measurements which will not be reported here. The EEG was recorded at a sampling rate of 250 Hz using Cz as the reference. All electrode impedances were maintained below 50 kOhm. The EGI system incorporates analog elliptical high- and low-pass with cut-off frequencies at 0.1 and 125 Hz (the Nyquist frequency), respectively. Filtering was performed before analog-to-digital conversion of the EEG.

Offline, the EEG data were analyzed using customized MATLAB scripts (R2011b, MathWorks Inc.) and the Fieldtrip toolbox (Oostenveld et al., 2011). Trials with response times longer than 2 s were excluded from all further analyses. The data were divided into epochs of sufficient length (–5 to +11 s around the onset of the first digit/flanking noise) to avoid data loss at the edges of the time-frequency representations due to windowing effects. The epoched data were bandpass filtered using an acausal sixth order IIR Butterworth filter between 0.5 and 45 Hz and re-referenced to the average of both mastoids. Before further analyses, 18 electrodes used for recording the electrooculogram (EOG) or positioned on the cheeks and jaw were removed for technical reasons.

Individual channels containing artifacts were identified through visual inspection and repaired by averaging over adjacent electrodes (according to the nearest neighbor approach implemented in the *ft_channelrepair* function in Fieldtrip). Data from one participant from the mild HL group were excluded from all further analyses due to a high number of artifact-contaminated channels. To remove further artifacts, an independent component analysis (ICA) was performed, and components containing eye blinks, saccadic eye movements, muscle activity, and heartbeats were identified by inspection of components' topographies and time courses and rejected. On average, 22% (SD: 6%) of the components were removed.

The time-frequency representation of oscillatory power in each trial was obtained by convolution of single trial time domain data with a family of Morlet wavelets (width: seven cycles). This analysis was performed for frequencies from 0.5 to 30 Hz in steps of 0.5 Hz and from –5 to +11 s around the onset of the first digit/flanking noise in steps of 0.05 s. Note that this long time interval included

the baseline period, encoding, delay, and probe (**Figure 1A**). The power of each time–frequency–electrode bin was calculated for each trial by taking the square norm of the complex wavelet coefficients. Adjustment for inter- and intra-individual variability in oscillatory power was performed by means of subtraction and division by the average power of the first 0.4 to 1 s of the baseline interval (relative change from baseline). For further analyses, each trial was split into the following periods: encoding, 0.4–4.8 s relative to first digit/flanking noise onset; delay, 0.4–1 s relative to the offset of the last digit/flanking noise; and probe, 0.4–1 s relative to probe-digit onset. All time intervals disregard the first 0.4 s as to not include evoked activity after stimulus on- or offset in the analysis.

STATISTICAL ANALYSES

A main motivation of the present study was to investigate the effect of HL on behavioral performance and alpha oscillations in the auditory Sternberg task. However, HL was confounded by age, as evidenced by a positive Pearson's correlation between age and PTA ($r = 0.44$, $p = 0.018$). To obtain a measure of HL that was independent of age, we calculated the residualized PTA, quantifying the variation in PTA across participants that could not be explained by age. In detail, the residualized PTA was estimated as the residuals of the linear regression of PTA on age. For the remainder of this paper, we will refer to the *z*-scored residualized PTA as 'rPTA.' In all further analyses, rPTA was included as a continuous covariate. Moreover, we considered it likely that brain compensatory mechanisms involved in overcoming the adverse listening conditions would not increase linearly with HL, but drop with more severe HL, especially under high memory load/background noise (see Introduction). To model this negative quadratic (inverted u-shape) relationship between HL and behavioral and brain responses, we additionally included the quadratic term rPTA-squared as a second continuous covariate in all further analyses.

Statistical analysis of behavioral data

First, we analyzed to what extent the individual adjustments of SNR levels were dependent on participants' HL. To evaluate whether individualization was needed, we calculated the Pearson's correlation between the 0 dB SRT80 value from the HINT and the non-residualized PTA.

In the auditory Sternberg task, response times were measured from the onset of the probe digit until the button press by the participant to indicate whether the probe digit appeared in the encoding. Accuracy was calculated as the percentage of correctly answered trials. Changes in task accuracy and response times as a function of the within-subject factors (memory load and background noise level) and the continuous between-subjects covariates (rPTA and rPTA-squared), were investigated using two separate repeated-measures ANCOVAs. All ANCOVAs showed violation of the assumption of sphericity (Mauchly's test, all $p < 0.05$), hence the Greenhouse–Geisser corrected *p*-values were calculated and reported for all results. Fisher's LSD tests were used for all *post hoc* analyses.

To illustrate the quadratic relationship between rPTA and response times (**Figure 2C**), a quadratic function was fitted to the response time as a function of rPTA using the least-squares approach implemented in the MATLAB functions *polyfit* and *polyval*.

Statistical analysis of EEG data

In the analysis of the EEG data, alpha power was averaged across frequencies from 6–12 Hz in a subset of 31 electrodes (**Figure 3A**, topographic maps) and across three time intervals outlined in **Figure 3A**: encoding, 0.4–4.8 s relative to the onset of the first digit/flanking noise; delay, 0.4–1 s relative to the offset of the last digit/flanking noise; and probe, 0.4–1 s relative to the onset of the probe digit. The 31 electrodes were chosen to derive a centro-parietal scalp distribution, which has previously been identified as an important site for alpha activity generation

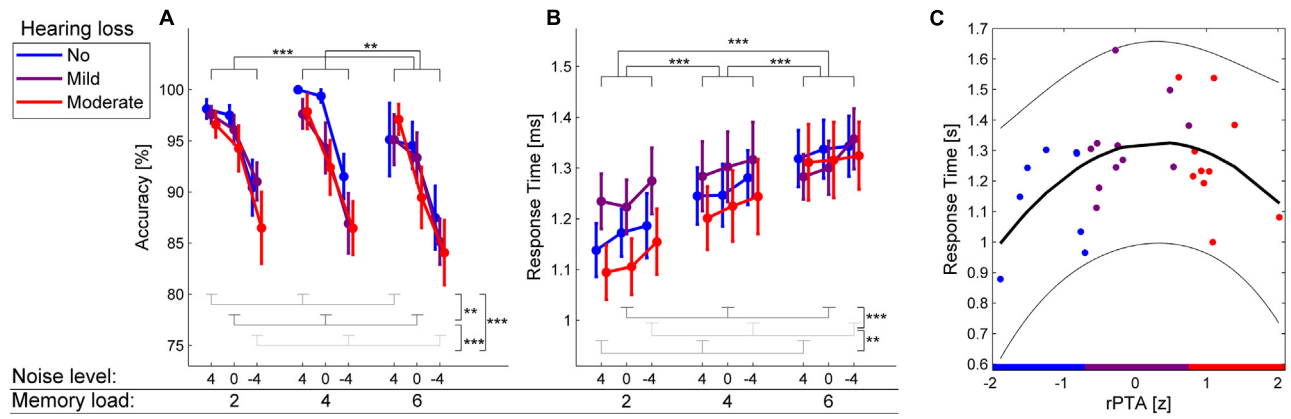


FIGURE 2 | Behavioral results. (A,B) Accuracy and response times in the auditory Sternberg task for participants with no HL (blue), mild HL (purple), and moderate HL (red) as a function of memory load (2, 4, 6 to-be-remembered items) and background noise level (4, 0, –4 dB SRT80). Error bars show ±1SEM. $^{**}p < 0.01$, $^{***}p < 0.001$. **(C)** Statistically significant quadratic regression between the *z*-scored rPTA and response times ($p = 0.025$). The least-squares regression line is shown in black. The 95% confidence interval is shown in thin lines. The slight overlap in rPTA of the three groups of HL is because the three groups were created before the impact of age on HL was regressed out (see Materials and Methods for details). Note that higher rPTA values indicate more severe HL.

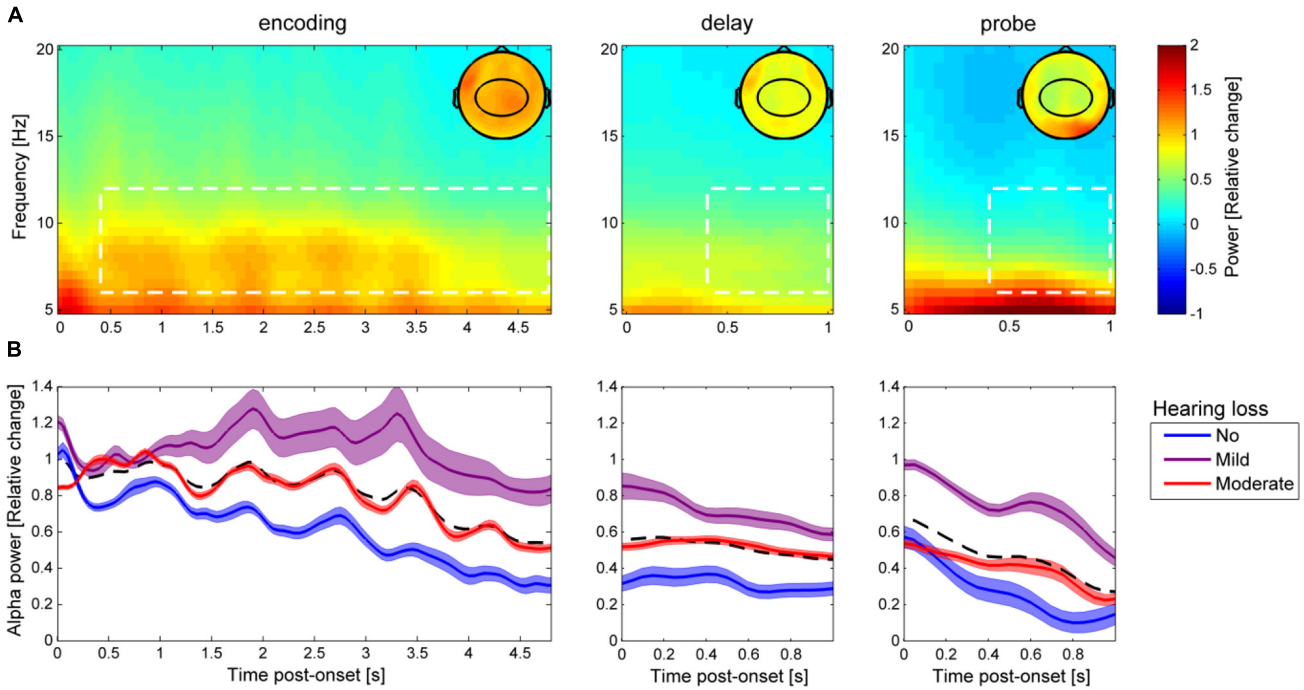


FIGURE 3 | Alpha power dynamics during the auditory Sternberg task. **(A)** Grand-average time-frequency power representation during encoding, delay, and probe (averaged across all participants, in all nine experimental conditions, and for all 31 centro-parietal electrodes highlighted in the topographic maps). The topographic maps show the spatial distribution of alpha power (6–12 Hz) averaged over the time-frequency data highlighted in the white dashed boxes, which were used for statistical analyses. **(B)** The bold lines show average alpha power in the three time periods (encoding, delay, and probe) separately for the three groups of HL (blue: no HL, purple: mild HL, red: moderate HL), with the colored areas indicating ±1 SEM. The black dashed line indicates the average over the three HL groups.

during auditory processing (Krause et al., 1996). Average alpha power during encoding, delay, and probe were subjected to three repeated-measures ANCOVAs with memory load and background noise level as within-subject factors and with rPTA and rPTA-squared as continuous between-subject covariates. All ANCOVAs showed violation of the assumption of sphericity (Mauchly's test, all $p < 0.05$), hence the Greenhouse–Geisser corrected p-values were calculated and reported for all results. All statistical analyses were performed using Statistica (version 12, StatSoft, Tulsa, OK, USA).

To illustrate the quadratic relationship between rPTA and alpha power (**Figure 4B**), the fitting procedure described in the section above was applied.

Studies have previously shown an interaction between response time and alpha activity (Klimesch, 2005). Relations between alpha activity during the probe period and response time were therefore evaluated using Pearson's correlation.

RESULTS

INDIVIDUAL ADJUSTMENTS OF SNR LEVELS

The individual adjustments of SNR levels using the SRT80 measure resulted in an average 0 dB SRT80 value of 4.61 dB [standard error of the mean (SEM) = 0.86], meaning that participants on average required an SNR level of 4.61 dB to successfully repeat 80% of words from sentences presented in noise. The 0 dB SRT80 values correlated positively with participants' non-residualized PTA ($r = 0.76$; $p < 0.001$). This indicates that participants with more severe HL required a higher SNR level of stimulus materials.

MEMORY LOAD, BACKGROUND NOISE LEVEL, AND HEARING LOSS IMPACT PERFORMANCE

Figure 2A shows the average accuracy for the three levels of memory load (2, 4, 6 digits) and the three background noise levels (–4 dB SRT80, 0 dB SRT80, 4 dB SRT80) in the auditory Sternberg task. The main effect of memory load on accuracy was significant [$F(2,50) = 6.26$, $p = 0.005$]. *Post hoc* tests revealed significantly increased accuracy for two compared with six items ($p < 0.001$) and for four compared with six items ($p = 0.002$) but not for two compared with four items ($p = 0.718$). Additionally, the main effect of background noise level on accuracy was significant [$F(2,50) = 28.35$, $p < 0.001$], with the *post hoc* analysis showing a significant decrease in accuracy with increasing noise level (all $p < 0.01$). There were no significant main effects of rPTA [$F(1,25) = 1.86$, $p = 0.185$) or rPTA-squared [$F(1,25) = 1.94$, $p = 0.176$], indicating that the degree of HL by itself did not significantly impact task accuracy. None of the interactions between background noise level, memory load, rPTA, and rPTA-squared were significant (all $p > 0.195$).

Figure 2B shows the average response times for the three memory loads and background noise levels. The main effect of memory load on response times was significant [$F(2,50) = 24.73$, $p < 0.001$]. *Post hoc* tests revealed significantly longer response

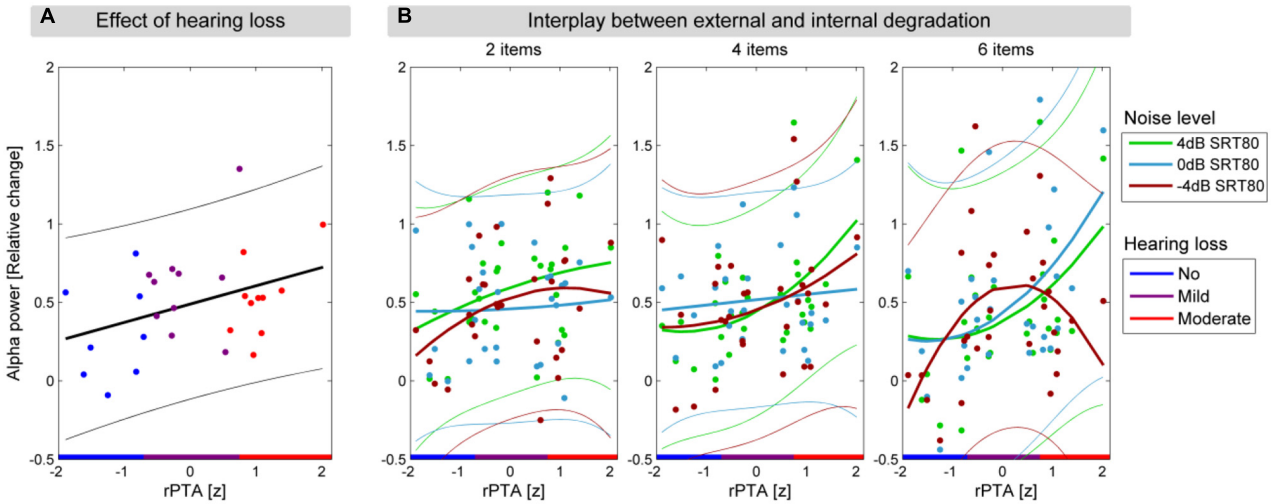


FIGURE 4 | Hearing loss affects alpha power in the delay period. **(A)** The significant linear relationship between alpha power in the delay interval and rPTA ($p = 0.048$). The regression line is shown with a solid black line, and the 95% confidence interval of the regression is shown in thin lines. **(B)** The three panels show the significant interaction between memory load, background noise level, and rPTA-squared, illustrated with quadratic fits between alpha power and rPTA for each background noise level (green: 4, light blue: 0, and dark red: −4 dB SRT80). Each panel shows one of the three memory load conditions (2, 4, and 6 items to be remembered) with alpha power during the delay interval as a function of rPTA with HL groups indicated on the *x*-axis (blue, no HL; purple, mild HL; red, moderate HL).

times for six compared with four and two to-be-retained digits, as well as for four compared with two digits (all $p < 0.001$). The main effect of background noise on response times was significant as well [$F(2,50) = 8.34$, $p = 0.001$]. *Post hoc* tests revealed significantly longer response times for the highest background noise level (–4 dB SRT80) compared with the intermediate noise level (0 dB SRT80; $p < 0.001$) and the lowest background noise level (4 dB SRT80; $p = 0.003$). Response times in the four and 0 dB SRT80 conditions did not differ significantly ($p = 0.328$). Interestingly, the main effect of rPTA-squared on response times was significant [$F(1,25) = 5.69$, $p = 0.025$]. This indicated a significant quadratic relationship between response times and the degree of HL in such a way that response times increased from no to mild HL, while response times decreased again for participants with the most severe HL (see **Figure 2C**). Neither the main effect of rPTA [$F(1,25) = 1.85$, $p = 0.185$], nor any interaction between memory load, background noise, rPTA, and rPTA-squared (all $p \geq 0.13$) reached significance.

TEMPORAL DYNAMICS OF ALPHA OSCILLATIONS

Figure 3A shows the grand-average baseline corrected time-frequency power representation (collapsed over all nine experimental conditions) for all participants throughout the encoding, delay, and probe periods of the auditory Sternberg task. The time course of alpha power (6–12 Hz; averaged over 31 scalp electrodes highlighted in topographic maps) for the three groups of HL are indicated in **Figure 3B**. Descriptively, alpha power decreased over the trial time course from encoding to delay and also during the probe interval. Normal hearing participants (no HL) exhibited the lowest alpha power in encoding, delay and probe, while the mild HL group showed the highest and the moderate HL group exhibited intermediate alpha power.

HEARING LOSS AFFECTS ALPHA OSCILLATIONS UNDER LOAD

We analyzed whether alpha power during the stimulus-free delay interval was dependent on memory load, background noise level, and HL. To this end, the average alpha power (6–12 Hz) across 31 centro-parietal electrodes during the delay interval (0.4–1 s relative to the offset of the background noise) was submitted to a repeated-measures ANCOVA with the factors memory load and background noise level and the continuous covariates rPTA and rPTA-squared. None of the main effects including background noise level [$F(2,50) = 1.23$, $p = 0.299$], memory load [$F(2,50) = 0.04$, $p = 0.598$], or rPTA-squared [$F(1,25) < 0.01$, $p = 0.989$] were significant. Importantly, however, the main effect rPTA was significant [$F(1,25) = 4.31$, $p = 0.0483$], indicating that alpha power during the delay increased significantly with the degree of HL (**Figure 4A**).

Moreover, the two-way interaction background noise level × rPTA-squared [$F(2,50) = 6.34$, $p = 0.004$] as well as the three-way interaction background noise level × rPTA-squared × memory load were significant [$F(4,100) = 2.86$, $p = 0.042$]. The direction of the significant three-way interaction is illustrated in **Figure 4B**. For the two lower memory loads (two and four to-be-remembered items), alpha power during the delay period increased moderately with the degree of HL for all background noise levels. This pattern of results changed significantly under the highest memory load (six to-be-remembered digits); here, alpha power strongly increased with HL under the two more favorable background noise levels (4 and 0 dB SRT80), but under the most severe background noise level (–4 dB SRT 80), alpha power increased only for participants with mild HL, whereas it decreased again for participants with moderate HL. The significant interaction between background noise level and rPTA ($p = 0.004$) is not shown, but resembles the same behavior

as observed for six items to be remembered shown in **Figure 4B**. None of the remaining interactions among rPTA, rPTA-squared, memory load, and background noise level were significant (all $p > 0.15$).

The main hypothesis of this experiment was focused on identifying condition and HL effects on alpha power during the delay. However, Obleser et al. (2012) also report smaller condition effects during the encoding and probe period. We therefore investigated alpha power during the encoding (0.4–4.8 s relative to the onset of the first digit/flanking noise) and probe (0.4–1 s relative to probe digit onset) interval as well. For the encoding interval, none of the main effects of memory load, background noise level, rPTA, and rPTA-squared, nor any interactions reached significance (all $p > 0.14$). During the presentation of the probe, a main effect of rPTA-squared was found [$F(1,25) = 9.63$, $p = 0.004$], while no other main effects or interactions were significant (all $p > 0.12$). Notably, an effect of rPTA-squared is also observed on the response time and the relationship between alpha activity during the probe, and the response time was investigated. A Pearson's correlation showed a positive relationship ($r = 0.35$, $p = 0.068$) between alpha power during the probe and response times, meaning that participants with higher alpha power during the probe interval showed longer response times. A similar relationship was not observed between the alpha power during the delay period and the response times ($r = 0.15$, $p = 0.42$).

DISCUSSION

In this study, we tested whether HL in older participants had an impact on the neural mechanisms of WM under changing task demands implemented by varying degrees of memory load and background noise. Our main findings can be summarized as follows: first, irrespective of HL, increasing memory load and higher background noise levels led to performance decrements in the auditory Sternberg paradigm. Second, the effects of the increasing memory load and background noise level on alpha activity during the delay were co-determined by the degree of HL. That is, participants suffering from a higher degree of HL exhibited a breakdown in alpha activity with increasing task difficulty, which was not observed for the participants with mild or no HL. These findings show how an internal auditory degradation (i.e., HL) interacts with external acoustic challenges during adverse listening.

THE EFFECT OF RETAINING AUDITORY STIMULI

Effects of WM processing on alpha power have been often observed only during the retention of stimuli in both auditory (van Dijk et al., 2010; Obleser and Weisz, 2012; Obleser et al., 2012; Becker et al., 2013; Scharinger et al., 2014) and visual tasks (Jensen et al., 2002; Schack and Klimesch, 2002; Sander et al., 2012b). It was therefore not unexpected that modulations of alpha power in this study were also found in the delay period.

The linear main effect of rPTA on alpha power in the delay period (**Figure 4A**) showed that alpha power increases with more severe HL, independent of task difficulty. This linear effect occurs despite the quadratic tendency seen in **Figure 3B**. The linear relationship in **Figure 4A** arose from large individual differences in alpha power, especially in the mildly impaired group, and was also affected by the residualization performed to remove age effects: first, this dependence of alpha power on HL is observed during the retention of the to-be-remembered digits, where no active listening is involved. Second, all participants were wearing hearing aids to equalize audibility of the digits presented during the encoding across participants. Interpreting the alpha activity as a sign of WM involvement (Jensen et al., 2002), our study shows that a higher degree of WM involvement is needed to overcome more severe HL to successfully retain the auditory information. This view of increased WM involvement with increased HL has been put forward in a number of studies (Pichora-Fuller and Singh, 2006; Rönnberg et al., 2008; Shinn-Cunningham and Best, 2008). The Ease of Language Understanding (ELU) model developed by Rönnberg et al. (2008) explains the involvement of the WM in speech understanding under adverse conditions. In detail, the ELU model builds on the ability to match auditory stimuli with a preexisting long-term memory store of phonological representations. When suffering from a HL, this match cannot readily be made due to the internal degradation. Hence WM processes are required for extracting acoustical cues that can trigger a phonological match and ensure a successful understanding. In line with the ELU model, the linear relationship between HL and alpha power can be interpreted as the increased WM resources needed to perform successful phonological matching in listeners with HL. Interestingly, the effect of HL on alpha activity is observed for participants wearing hearing aids, which is thought to ensure equal audibility, but arguably cannot restore the WM resources needed to retain speech stimuli.

Hearing aids can indeed ensure audibility and restore intelligibility in quiet situations, while other aspects of listening, such as processing of temporal cues, are not alleviated by amplification (Ardoint et al., 2010). Furthermore, speech intelligibility in noisy situations also remains affected by HL and cannot be fully restored by amplification (Plomp, 1978; Dillon, 2001). This is indeed evident from the positive relationship between HL and the 0 dB SRT80 value. Peelle et al. (2011) found that increased HL was correlated with decreased gray matter volume of the auditory cortex, i.e., a structural change in the brain. If HL causes structural changes in the auditory cortex, this might explain why individual HL compensation via amplification does not nullify such structural deviation in the auditory system, and HL-dependent effects, such as the present ones, are observed despite hearing aids being employed.

The impact of the experimental conditions (memory load and background noise level) proved only to be significant in interactions with HL. Our results showed that when increasing the external degradation, i.e., the background noise level, an increase in alpha activity with HL was observed for the lower levels of background noise. However, for the highest background noise level, a breakdown in alpha activity was observed for the participants with the most severe degree of HL tested in this study (moderate HL). This breakdown in alpha power is only observed when participants have to remember six digits in the most difficult noise condition (**Figure 4B**). The almost linear increase in alpha power with HL severity observed at lower background noise levels (4 and 0 dB SRT80) suggests that although the noise levels

are individualized, participants with increased HL require additional WM resources to be able to perform the task. Indeed, it has previously been suggested that people suffering from HL need to allocate additional resources to process auditory information (Rabbitt, 1991). The findings in this study lend neural support to this hypothesis.

The breakdown in alpha power with increased HL and background noise level further suggests that the participants suffering from reduced hearing reach a ceiling at which no further enhancement in alpha activity can be achieved, and alpha power begins to decrease. Such alpha power breakdown has been observed before when older participants, not considering HL, are subjected to a higher WM load in a visual Sternberg task, while no effect of age was observed on task accuracy (Sander et al., 2012b). Similar findings of neural activity breakdown with high WM loads for increasing age have been observed in fMRI studies (Reuter-Lorenz and Cappell, 2008; Schneider-Garces et al., 2010; Grady, 2012). Also here, the activity breakdown is not necessarily accompanied by changes in task accuracy. According to the "compensation-related utilization of neural circuits" (CRUNCH) hypothesis, the brain increases its activation to engage more neural resources as a result of aging, independent of WM involvement. However, with increasing WM demands, this recruitment reaches a ceiling, and the activity decreases, although no changes in task performance are observed (Reuter-Lorenz and Cappell, 2008). We suggest that, similar to increasing age, more severe HL can cause neural activity breakdown as a result of having to engage more WM resources than participants with better hearing. It is believed that the cause of the observed breakdown is a combination of the two observations that: participants with more severe HL experience generally higher WM involvement (independent of experimental conditions, **Figure 4A**) and during WM tasks they have increased WM involvement (**Figure 4B**). To our knowledge our results are the first to demonstrate a breakdown of neural activity with increased HL.

Alpha power during the delay was affected by memory load in a three-way interaction with background noise and rPTA-squared. Our experimental design was modified from the auditory Sternberg task applied by Obleser et al. (2012), who found main effects of both memory load and auditory degradation (obtained through noise-vocoding of the digits) on alpha activity. The lack of a main effect of memory load in the present study might be explained best by the differences in participants (older hearing impaired vs. younger normal hearing), rather than auditory degradation (background noise vs. noise-vocoding). Both of these changes were introduced to achieve some gain in external validity in the present study.

Although we corrected for the difference in age between participants in this study, we cannot account for the average differences between younger and older persons, which has been proven to affect both alpha activity and WM resources (Klimesch, 1999; Sander et al., 2012a). Although increased age might have resulted in participants having generally less WM resources available and thereby reaching alpha power breakdown, differences in cohort age between the studies cannot explain the non-significant main effect of memory load in the present study. We suggest that the lack of memory load effect can be explained by the fact that the hearing impaired participants are already performing at ceiling and cannot further increase their alpha activity when subjected to higher memory loads and/or background noise levels. This statement is supported by two observations: firstly, that the alpha power increased with HL, independent of the experimental condition. Secondly, that the conditions effects (rPTA × background noise level and rPTA × background noise level × memory load) showed a decrease in alpha power for the moderately impaired participants, c.f. **Figure 4B**.

NO EFFECTS OF HEARING LOSS ON TASK ACCURACY

To adjust for the differences in HL, the background noise levels were individualized using the SRT80 measure obtained from the HINT test (for details see Materials and Methods). The positive relation between HL and 0 dB SRT80 shows that for participants with more severe HL a lower background noise level (i.e., higher 0 dB SRT80) is needed. This relationship emphasizes the importance of individualizing the background noise level to ensure equal task accuracy across all participants, independent of HL. Indeed, the non-significant effect of HL on task accuracy confirms the success of applying individual noise levels.

As hypothesized, task accuracy significantly decreased both with increased memory load and background noise level. As **Figure 2A** shows, background noise levels showed stronger effects on task accuracy than changes in the memory load. In line with the modulations of alpha activity, this finding emphasizes that auditory degradation induces a larger WM involvement than changes in the memory load for the memory loads and background noise level tested in this study. Significant effects of the experimental conditions on task accuracy have sometimes been reported in auditory and visual Sternberg tasks (Rojas et al., 2000; Jensen et al., 2002; Sander et al., 2012b), but most studies aim at having no condition effects on accuracy (Sternberg, 1966; Lehtelä et al., 1997; Leiberg et al., 2006; Obleser et al., 2012). As noted by Rojas et al. (2000), the confounding effect of task accuracy on response time and alpha activity makes it impossible to determine whether WM processing is indeed involved in solving the task, especially for wrongly answered trials. In this study, effects of memory load and background noise level on task accuracy were found, which is a limitation of the study. However, obtaining task accuracies close to 100% correct for all conditions and participants would require troublesome and time consuming individualization. Alternatively, including only the correctly answered trials in the current analysis would result in an unfeasibly low number of trials per condition. However, as we observe effects of HL on the alpha power, we believe that WM processing was involved during task solving.

The response times were affected both by the experimental conditions (**Figure 2B**) and HL (**Figure 2C**), the latter showing a speed-up in response times with increased HL. As a sign of stimulus retrieval (Sternberg, 1966), it was expected that the response time would show effects of the experimental condition as well as HL. The increase in response times from normal to mildly impaired hearing suggests that increasing internal degradation of the auditory signal results in longer processing times of the probe digit. As HL increases from mild to moderate, participants' strategy might change resulting in shorter response times (**Figure 2C**).

The effect of rPTA-squared on alpha activity during the probe also proved to be significant and although the correlation between the alpha activity during the probe and the response times only approached significance ($p = 0.068$), we believe that the changes in alpha power during the probe period arise from changes in the speed of information processing (Klimesch, 2005) and not WM processing as such.

In summary, the present findings suggest that despite being compensated for the loss of hearing through hearing aid amplification and by individually setting the administered signal-to-noise ratios, higher degrees of HL are detrimentally affecting a cardinal neural mechanism of overcoming adverse listening conditions, namely the increase in posterior alpha power. Apparently, participants with moderate HL reach a ceiling level at which no more WM resources can be recruited, and thus alpha power begins to decrease again. These findings not only reveal that hearing aid amplification by itself is not sufficient for restoring normal neural signatures of auditory processing, but also suggest that persons suffering from a higher degree of HL reach a WM limit at a lower task demand.

ACKNOWLEDGMENTS

EBP is supported by a grant from the Oticon Foundation. We wish to thank all the participants in the experiment and Gunilla Wänström, Irene Slättengren, and Mathias Hällgren for their assistance during the experiment. The authors are grateful for the helpful discussions with researchers at Eriksholm Research Centre and members of the Max Planck Research Group "Auditory Cognition."

REFERENCES

Ardoint, M., Sheft, S., Fleuriot, P., Garnier, S., and Lorenzi, C. (2010). Perception of temporal fine-structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss. *Int. J. Audiol.* 49, 823–831. doi: 10.3109/14992027.2010.492402

Baddeley, A. (1986). *Working Memory*. Oxford: Oxford University Press.

Becker, R., Pefkou, M., Michel, C. M., and Hervais-Adelman, A. G. (2013). Left temporal alpha-band activity reflects single word intelligibility. *Front. Syst. Neurosci.* 7:121. doi: 10.3389/fnsys.2013.00121

Bonnefond, M., and Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Curr. Biol.* 22, 1969–1974. doi: 10.1016/j.cub.2012.08.029

CHABA. (1988). Speech understanding and aging. *J. Acoust. Soc. Am.* 83, 859–895. doi: 10.1121/1.395965

Dillon, H. (2001). Hearing aids. *Gac. Med. Mex.* 126, 417–418. doi: 10.1007/SpringerReference_115704

Drullman, R., Wagener, K., Wouters, J., and Eneman, K. (2005). *Speech Recognition Tests for Different Languages.* Report FP6-004171. Hearing in the communication society. Available at: www.hearcom.eu

Grady, C. (2012). The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* 13, 491–505. doi: 10.1038/nrn3256

Hällgren, M., Larsby, B., and Arlinger, S. (2006). A swedish version of the Hearing In Noise Test (HINT) for measurement of speech recognition. *Int. J. Audiol.* 45, 227–37. doi: 10.1080/14992020500429583

Jensen, O., Gelfand, J., Kounios, J., and Lisman, J. E. (2002). Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* 12, 877–882. doi: 10.1093/cercor/12.8.877

Jensen, O., and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* 4:186. doi: 10.3389/fnhum.2010.00186

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–95. doi: 10.1016/S0165-0173(98)00056-3

Klimesch, W. (2005). "EEG Alpha and Cognitive Processes," *in Time and the Brain,* ed. R. Miller (Amsterdam: Harwood Academic Publishers), 236–261.

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/j.brainresrev.2006.06.003

Krause, C. M., Lang, A. H., Laine, M., Kuusisto, M., and Pörn, B. (1996). Event-related EEG desynchronization and synchronization during an auditory memory task. *Electroencephalogr. Clin. Neurophysiol.* 98, 319–326. doi: 10.1016/0013-4694(96)00283-0

Larsby, B., Hällgren, M., and Leandersson, M. (2011). "The swedish digit triplet test for screening of hearing by telephone and internet," *in Proceedings of the First International Conference on Cognitive Hearing Science for communication*, Linköping.

Lehtelä, L., Salmelin, R., and Hari, R. (1997). Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neurosci. Lett.* 222, 111–114. doi: 10.1016/S0304-3940(97)13361-4

Leiberg, S., Lutzenberger, W., and Kaiser, J. (2006). Effects of memory load on cortical oscillatory activity during auditory pattern working memory. *Brain Res.* 1120, 131–140. doi: 10.1016/j.brainres.2006.08.066

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42, S49–S58. doi: 10.3109/14992020309074624

Lunner, T., Rudner, M., and Rönnberg, J. (2009). Cognition and hearing aids. *Scand. J. Psychol.* 50, 395–403. doi: 10.1111/j.1467-9450.2009.00742.x

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. a., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol. A.* 58, 22–33. doi: 10.1080/02724980443000151

Obleser, J., and Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb. Cortex* 22, 2466–77. doi: 10.1093/cercor/bhr325

Obleser, J., Wöstmann, M., Hellbernd, N., Wilsch, A., and Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. *J. Neurosci.* 32, 12376–12383. doi: 10.1523/JNEUROSCI.4908-11.2012

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 156869. doi: 10.1155/2011/156869

Peelle, J. E., Troiani, V., Grossman, M., and Wingfield, A. (2011). Hearing loss in older adults affects neural systems supporting speech comprehension. *J. Neurosci.* 31, 12638–12643. doi: 10.1523/JNEUROSCI.2559-11.2011

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *Int. J. Audiol.* 42(Suppl. 2), 2S26–2S32.

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* 63, 533–549. doi: 10.1121/1.381753

Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol. Suppl.* 476, 167–176. doi: 10.3109/00016489109127274

Reuter-Lorenz, P. A., and Cappell, K. A. (2008). Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* 17, 177–182. doi: 10.1111/j.1467-8721.2008.00570.x

Rojas, D. C., Teale, P. D., Sheeder, J. L., and Reite, M. L. (2000). Neuromagnetic alpha suppression during an auditory sternberg task. Evidence for a serial, self-terminating search of short-term memory. *Brain Res. Cogn. Brain Res.* 10, 85–9. doi: 10.1016/S0926-6410(00)00026-4

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S99–S105. doi: 10.1080/14992020802301167

Sander, M. C., Lindenberger, U., and Werkle-Bergner, M. (2012a). Lifespan age differences in working memory: a two-component framework. *Neurosci. Biobehav. Rev.* 36, 2007–2033. doi: 10.1016/j.neubiorev.2012.06.004

Sander, M. C., Werkle-Bergner, M., and Lindenberger, U. (2012b). Amplitude modulations and inter-trial phase stability of alpha-oscillations differentially reflect working memory constraints across the lifespan. *Neuroimage* 59, 646–54. doi: 10.1016/j.neuroimage.2011.06.092

Schack, B., and Klimesch, W. (2002). Frequency characteristics of evoked and oscillatory electroencephalic activity in a human memory scanning task. *Neurosci. Lett.* 331, 107–110. doi: 10.1016/S0304-3940(02)00846-7

Scharinger, M., Herrmann, B., Nierhaus, T., and Obleser, J. (2014). Simultaneous EEG-fMRI brain signatures of auditory cue utilization. *Front. Neurosci.* 8:137. doi: 10.3389/fnins.2014.00137

Schneider-Garces, N. J., Gordon, B. a., Brumback-Peltz, C. R., Shin, E., Lee, Y., Sutton, B. P., et al. (2010). Span, CRUNCH, and beyond: working memory capacity and the aging brain. *J. Cogn. Neurosci.* 22, 655–669. doi: 10.1162/jocn.2009.21230

Shinn-Cunningham, B. G., and Best, V. (2008). Selective attention in normal and impaired hearing. *Trends Amplif.* 12, 283–299. doi: 10.1177/1084713808325306

Simonsen, C. S., and Behrens, T. (2009). A new compression strategy based on a guided level estimator. *Hear. Rev.* 16, 26–31.

Stenfelt, S., and Rönnberg, J. (2009). The signal-cognition interface: interactions between degraded auditory signals and cognitive processes. *Scand. J. Psychol.* 50, 385–393. doi: 10.1111/j.1467-9450.2009.00748.x

Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654. doi: 10.1126/science.153.3736.652

Strauß, A., Wöstmann, M., and Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Front. Hum. Neurosci.* 8:350. doi: 10.3389/fnhum.2014.00350

van Dijk, H., Nieuwenhuis, I. L. C., and Jensen, O. (2010). Left temporal alpha band activity increases during working memory retention of pitches. *Eur. J. Neurosci.* 31, 1701–7. doi: 10.1111/j.1460-9568.2010.07227.x

Van Engen, K. J., and Peelle, J. E. (2014). Listening effort and accented speech. *Front. Hum. Neurosci.* 8:577. doi: 10.3389/fnhum.2014.00577

Wagener, K., Josvassen, J. L., and Ardenkjær, R. (2003). Design, optimization and evaluation of a Danish sentence test in noise. *Int. J. Audiol.* 42, 10–17. doi: 10.3109/14992020309056080

Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *J. Am. Acad. Audiol.* 17, 487–497. doi: 10.3766/jaaa.17.7.4

Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood. what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Wöstmann, M., Herrmann, B., Wilsch, A., and Obleser, J. (2015). Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *J. Neurosci.* 35, 1458–1467. doi: 10.1523/JNEURDSCI. 3250-14.2015

Conflict of Interest Statement: Eriksholm Research Centre is part of Oticon A/S and as such the salary of Eline Borch Petersen and Thomas Lunner were paid by Oticon A/S. Hearing aids were provided by Oticon A/S.

Received: 02 October 2014; accepted: 04 February 2015; published online: 19 February 2015.

Citation: Petersen EB, Wöstmann M, Obleser J, Stenfelt S and Lunner T (2015) Hearing loss impacts neural alpha oscillations under adverse listening conditions. Front. Psychol. ***6****:177. doi: 10.3389/fpsyg.2015.00177*

This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology.

Copyright © 2015 Petersen, Wöstmann, Obleser, Stenfelt and Lunner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 May 2015
doi: 10.3389/fpsyg.2015.00597

Cognitive spare capacity: evaluation data and its association with comprehension of dynamic conversations

Gitte Keidser[1], Virginia Best[2], Katrina Freeston[1] and Alexandra Boyce[3]*

[1] National Acoustic Laboratories, Sydney, NSW, Australia, [2] Department of Speech, Language, and Hearing Sciences, Boston University, Boston, MA, USA, [3] Department of Audiology, Macquarie University, Sydney, NSW, Australia

It is well-established that communication involves the working memory system, which becomes increasingly engaged in understanding speech as the input signal degrades. The more resources allocated to recovering a degraded input signal, the fewer resources, referred to as cognitive spare capacity (CSC), remain for higher-level processing of speech. Using simulated natural listening environments, the aims of this paper were to (1) evaluate an English version of a recently introduced auditory test to measure CSC that targets the updating process of the executive function, (2) investigate if the test predicts speech comprehension better than the reading span test (RST) commonly used to measure working memory capacity, and (3) determine if the test is sensitive to increasing the number of attended locations during listening. In Experiment I, the CSC test was presented using a male and a female talker, in quiet and in spatially separated babble- and cafeteria-noises, in an audio-only and in an audio-visual mode. Data collected on 21 listeners with normal and impaired hearing confirmed that the English version of the CSC test is sensitive to population group, noise condition, and clarity of speech, but not presentation modality. In Experiment II, performance by 27 normal-hearing listeners on a novel speech comprehension test presented in noise was significantly associated with working memory capacity, but not with CSC. Moreover, this group showed no significant difference in CSC as the number of talker locations in the test increased. There was no consistent association between the CSC test and the RST. It is recommended that future studies investigate the psychometric properties of the CSC test, and examine its sensitivity to the complexity of the listening environment in participants with both normal and impaired hearing.

Keywords: cognitive spare capacity, working memory capacity, updating, speech comprehension, dynamic speech test

OPEN ACCESS

Edited by:
Carine Signoret, Linnaeus Centre for Hearing and Deafness, Sweden

Reviewed by:
Patrik Sörqvist, University of Gävle, Sweden
Sushmit Mishra, Utkal University, India

****Correspondence:***
Gitte Keidser, National Acoustic Laboratories, Australian Hearing Hub, 16 University Avenue, Macquarie University, Sydney, NSW 2109, Australia
gitte.keidser@nal.gov.au

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *12 February 2015*
Accepted: *22 April 2015*
Published: *06 May 2015*

Citation:
Keidser G, Best V, Freeston K and Boyce A (2015) Cognitive spare capacity: evaluation data and its association with comprehension of dynamic conversations. Front. Psychol. 6:597. doi: 10.3389/fpsyg.2015.00597

Abbreviations: 4FA HL, four-frequency average hearing loss; ANOVA, analysis of variance; CSC, cognitive spare capacity; CSCT, cognitive spare capacity test; ILTASS, international long-term average speech spectrum; RST, reading span test; SE, standard error; SNR, signal-to-noise ratio; SRT, speech reception threshold.

Introduction

Participation in social activities has been found to be important for a person's psychological and general well-being (Pinquart and Sörensen, 2000), and verbal communication is often the key to social interactions. Effective communication requires an interaction between implicit bottom–up and explicit top–down processes, and thus relies on both healthy auditory and cognitive systems (Wingfield et al., 2005; Pichora-Fuller and Singh, 2006; Schneider et al., 2010). Higher-level processing of speech, such as comprehension, inference making, gist formulation, and response preparation, involves in particular working memory processing (Daneman and Carpenter, 1980; Schneider et al., 2007; Wingfield and Tun, 2007). Working memory is defined as a limited capacity system with storage and processing capabilities that enables the individual to temporarily hold and manipulate information in active use as is necessary for comprehending speech (Baddeley, 1992; Just and Carpenter, 1992). In the widely accepted multi-component model of working memory, first introduced by Baddeley and Hitch (1974), the central executive is considered the control system for manipulation of input to either the phonological loop, visuospatial sketchpad, or episodic buffer (Repovš and Baddeley, 2006), and is considered the component that most influences working memory processing efficiency (McCabe et al., 2010). According to Miyake et al. (2000), the executive function is associated with three organizational processes; inhibition, shifting, and updating. When related to speech comprehension, these three processes refer to the ability to ignore irrelevant information, select the conversation to follow, and process the most recent sounds in order to compare items with stored knowledge to infer meaning, respectively.

Several speech perception models have been proposed to more specifically explain the mechanism of speech comprehension from sensory information, such as the cohort (Marslen-Wilson and Tyler, 1980; Marslen-Wilson, 1990), TRACE (McClelland and Elman, 1986; McClelland, 1991), and neighborhood activation (Luce and Pisoni, 1998) models. A more recent addition is the ease of language understanding (ELU) model (Rönnberg et al., 2008, 2013) that differs from the earlier models by its assumption that explicit working memory capacity is called for whenever there is a mismatch between the input signal and the phonological representations in long-term memory (Rönnberg et al., 2013). In brief, the ELU model stipulates the interaction between an implicit processing path and a slower explicit processing loop that run in parallel. While the multimodal input signal matches a sufficient number of phonological attributes in the mental lexicon, the lexical access proceeds rapidly and automatically along the implicit processing path with little engagement of the explicit processing loop. The explicit processing loop, which uses both phonological and semantic long-term memory information to attempt to understand the gist of the conversation, is, however, increasingly accessed when there is a mismatch between input signal and the phonological representations in long-term memory.

According to the ELU model, explicit working memory processing, including the executive processes, is increasingly relied on to infer meaning as the input signal becomes less clear and the listening situation more challenging. This notion is supported by several studies, which have shown that people with higher working memory capacity are less susceptible to distortion introduced by such factors as hearing impairment, increased complexity in the environment, or the introduction of unfamiliar signal processing in hearing devices; i.e., are better at understanding speech under such conditions (Lunner, 2003; Lyxell et al., 2003; Rudner et al., 2011a; Arehart et al., 2013; Meister et al., 2013). In these studies, a dual-task test, known as the RST (Daneman and Carpenter, 1980; Rönnberg et al., 1989), was used to measure the combined storage and processing capacity of working memory. The RST presents participants with a written set of unrelated and syntactically plausible sentences. After each sentence participants have to indicate if the sentence was sensible (e.g., the boy kicked the ball) or not (e.g., the train sang a song), and after a span of sentences they have to recall either the first or last word in the sentences (ignoring the article). Participants are presented with an increasingly longer span of sentences from 3 to 6. Performance on this paradigm has been found to be well-associated with speech comprehension (Daneman and Merikle, 1996; Akeroyd, 2008), and thus seems to be a solid predictor of inter-individual differences in speech processing abilities.

Recently, there has been an increased interest in the audiological community to prove that intervention with hearing devices, or specific device features, reduces cognitive resources allocated to listening; i.e., frees up resources for other cognitive processes such as higher-level speech processes (Sarampalis et al., 2009; Ng et al., 2013). This calls for an auditory test that taps into the cognitive functions engaged when communicating, such as working memory and the executive processes, and that is sensitive to different types of distortion and so can measure intra-individual differences in cognitive listening effort as the quality of the input changes. As one example of such a test, the concept of the RST was applied to the Revised Speech in Noise test to specifically investigate working memory capacity for listening to speech in noise (Pichora-Fuller et al., 1995). Using a mixture of high- and low-context sentences, participants were presented with a span of sentences and asked at the end of each sentence to indicate whether the final word was predictable from the sentence context or not, and at the end of the span to recall the final words. The authors found that age and increasing background noise disturbed the encoding of heard words into working memory, reducing the number of words that could be recalled.

New paradigms have also been introduced that aim to measure the CSC, defined as the residual capacity available for processing heard information after successful listening has taken place (Rudner et al., 2011b). An example is the CSCT, introduced by Mishra et al. (2013a), that taps into an individual's working memory storage capacity, multimodal binding capacity (when visual cues are present), and executive skills after resources have been used for processing the heard stimuli. In this test participants are presented with lists of two-digit numbers, spoken randomly by a male or female talker, and are either asked to recall the highest (or lowest) numbers spoken by each talker, or to recall the

odd (or even) numbers spoken by a particular talker. Thus the test measures the ability to update or inhibit information, respectively, and then recall the information, after resources have been spent on recognizing what has been said. The authors have argued that CSC as measured with the CSCT is different from general working memory capacity as measured with the RST. This is a reasonable assumption when considering the overall mental processes involved in the two tests. For example, the RST requires intake of written sentences, analysis of semantic content, formulation and delivery of a response, and storage and recall of words, whereas the CSCT requires attention to and processing of heard stimuli (potentially degraded by some form of distortion), a decision to be made about what to store, and storage, deletion, and recall of numbers. While there is some overlap in processes, there are also substantial differences, and therefore one would not expect a perfect correlation between performances on the two tests. Further, while reading the sentences in the RST for most people would be an implicit process, listening to the stimuli in the CSCT may require explicit processing as stipulated by the ELU model. That is, the CSC would be expected to be increasingly reduced under increasingly demanding listening conditions where explicit resources become involved in the processes of recognizing the input signal, leaving fewer resources for completing the remaining operations required by the CSCT. Therefore, it is likely that the residual capacity measured with CSCT under adverse test conditions is something less than the full working memory capacity measured with the RST. The authors of the CSCT have further suggested that during the updating or inhibition process of CSCT, if an executive resource that is required for performing these tasks has been depleted in the process of recognizing the numbers, the function of this particular resource may be at least partially compensated for by another cognitive resource that is separate from working memory. Consequently, a measure of working memory capacity may not adequately assess CSC. The CSCT has been evaluated with normal-hearing and hearing-impaired listeners under different conditions (Mishra et al., 2013a,b, 2014). Overall, the results, which are presented in more detail in the next section, suggested that the test has merit as a measure of cognitive listening effort. In addition, there was no overall association between CSCT and RST scores, suggesting that CSCT is not merely a measure of working memory capacity. In this paper we present an English version of the CSCT.

A hypothesis that a measure of CSC would better predict communicative performance than a measure of working memory capacity as captured with the RST (Mishra et al., 2013a) has not been investigated. Thus, we investigate in this paper if the CSCT or RST better predicts speech comprehension in noise. We recently developed and introduced a speech comprehension test that is designed to more closely resemble real world communication (Best et al., in review). This paradigm has been extended to include monologs and dialogs between 2 and 3 spatially separated talkers to study dynamic aspects of real communication. As the CSCT is designed to be administered under conditions similar to those in which speech performance is measured, it seems to provide an excellent tool for objectively investigating the cognitive effect of changing complexity of the listening conditions within individuals. We, therefore, further use the CSCT to investigate if dynamic changes in voice and location like those in our new speech test affect listening effort, as reflected in CSC.

In summary, this paper presents two experiments to address three aims. The aim of the first experiment is to present and evaluate an English version of the CSCT. The aims of the second experiment are to examine if CSC is a better predictor than working memory capacity of speech comprehension in noise, and to examine if increasing the number of talkers in the listening situation reduces CSC. In both experiments, listening conditions were simulated to represent, as best as possible, realistic listening environments. Treatment of test participants was approved by the Australian Hearing Ethics Committee and conformed in all respects to the Australian government's National Statement on Ethical Conduct in Human Research.

Experiment I

The aim of Experiment I was to evaluate an English version of the CSCT. The original Swedish test by Mishra et al. (2013a) was designed to measure both inhibition and updating. Different lists of thirteen two-digit numbers spoken randomly by a male and a female talker were made up for each task. For either task the listener was asked to remember at least two items. In the inhibition task, listeners were asked to remember the odd or even number spoken by one of the talkers, meaning they had to inhibit numbers spoken by the non-target talker. In the updating task, the task was to remember the highest or lowest number spoken by each talker, meaning that the listener had to update information stored in working memory when a new number met the criterion. Each list was designed to present three or four inhibition or updating events. A high memory load condition was created in which the listeners were further asked to remember the first number of the list, although this number was not taken into account in the final score.

In three studies, the Swedish version of the CSCT was evaluated by studying sensitivity to memory load (low vs. high), noise (quiet vs. stationary speech-weighted noise vs. modulated speech-like noise), and presentation modality (audio vs. audio-visual) in young normal-hearing and older hearing-impaired listeners (Mishra et al., 2013a,b, 2014). The older hearing-impaired listeners had stimuli amplified to compensate for their hearing loss, and for the noise conditions the SNR were individually selected to ∼90% recognition in the stationary noise. Overall, the studies showed that the older hearing-impaired listeners generally had reduced CSC relative to the younger normal-hearing listeners. For both populations, increasing the memory load and listening in stationary noise relative to quiet reduced CSC. Relative to quiet, the highly modulated speech-like noise reduced CSC in the older, but not in the younger cohort. The older hearing-impaired listeners also showed reduced CSC when listening in audio-only mode relative to audio-visual mode in noise and in quiet. Relative to the audio-visual mode, the younger normal-hearing listeners showed reduced CSC in audio-only mode when listening in noise, but increased CSC when listening in quiet. The authors argued

that in all cases where CSC was relatively reduced, more pressures were put on the available cognitive resources needed for the act of listening, and that in the more demanding listening conditions visual cues counteracted for the disruptive effect of noise and/or poorer hearing (Mishra et al., 2013a,b, 2014).

In the studies conducted by Mishra et al. (2013a,b, 2014), task never interacted with any of the other factors, suggesting that the inhibition and updating measures were equally sensitive to different changes in the test condition. This is presumably because inhibition can be considered a part of the updating task, as items needed to be suppressed from working memory when a new item that fitted the criterion was stored. Consequently, to simplify the test design only the updating task was used in this study. The updating task was selected because the inhibition task in the Mishra studies generally produced higher scores than the updating task, with scores being close to ceiling for normal-hearing listeners. The decision to exclude the inhibition task meant that the need to switch between talker gender in the stimulus material was not strictly needed. There is a general belief that hearing-impaired people have more difficulty understanding female voices due to their more high-pitched characteristic (e.g., Helfer, 1995; Stelmachowicz et al., 2001), a factor that could have influenced the reduced CSC measured in the older hearing-impaired listeners by Mishra et al. (2014). To explore this further, we decided to present the updating task spoken by single talkers (one male or one female within each list), to test the effects of individual differences in talker characteristics (potentially including gender effects) on CSC. Removing the gender effect within lists meant that the listener did not have to attend to the talker gender during testing. On the other hand, the number of updating events in each list increased to four or five, with three lists introducing six updating events.

Like the Swedish version, the English version was further evaluated for sensitivity to population group (younger normal-hearing vs. older hearing-impaired listeners), noise (quiet vs. babble-noise vs. cafeteria noise), and presentation modality (audio only vs. audio-visual). While the Swedish test was evaluated under headphones with target and noise presented co-located, and in artificial noises, we chose to evaluate the CSCT under more natural listening conditions by presenting target and noise spatially separated in the free field, and using more realistic background noises. Introducing spatial separation in our presentation was expected to ease segregation (Helfer and Freyman, 2004; Arbogast et al., 2005), and hence the load on the executive function, for both normal-hearing and hearing-impaired listeners. However, this advantage was anticipated to be counteracted for during testing by choosing individual SNRs corresponding to the same speech recognition target used by Mishra et al. (2013b, 2014). Unlike the noises used by Mishra et al. (2013b, 2014) our babble- and cafeteria-noises were made up from intelligible discourses and conversations, respectively. As a result, our babble-noise was slightly more modulated than Mishra's stationary noise, whereas our cafeteria-noise was slightly less modulated than Mishra's speech-like noise. Finally, as in the Mishra studies, performance on the CSCT was related to measures of working memory capacity as measured with the RST and an independent test of updating. Overall, we expected to reproduce the findings by Mishra et al. (2014) with respect to the effect of population group, noise, and presentation modality, and we predicted that only the older hearing-impaired listeners would be affected by individual talker differences.

Methodology

Participants

Participants included 11 females and 10 males recruited among colleagues and friends of the authors. Among the 21 participants, 12 could be considered younger normal-hearing listeners. Their average age was 31.6 years (ranging from 22 to 49 years), and their average bilateral 4FA HL, as measured across 0.5, 1, 2, and 4 kHz, was 0.4 dB HL (SE = 1.0 dB). The average age of the remaining nine participants was 72.3 years (ranging from 67 to 77 years), and they presented an average 4FA HL of 29.9 dB HL (SE = 3.0 dB). This group is referred to as older hearing-impaired listeners, although it should be noted that the hearing losses were generally very mild with the greatest 4FA HL being 46.3 dB HL. Participants were paid a small gratuity for their inconvenience.

The Stimuli

The stimulus material to measure CSC for updating was adapted from Mishra et al. (2013a). Audio-visual recordings of two-digit numbers were obtained using one male and one female native English speaker with Australian accents narrating the numbers 11–99 sequentially. Recordings were performed in an anechoic chamber, with the talkers wearing dark clothes and seated in front of a gray screen. Video recordings, showing head and shoulders of the talkers, were obtained using a Legria HFG10 Canon video-camera set at 1920 × 1080 resolution. Three high-powered lights were positioned to the sides and slightly in front of the talker, facing away from them and reflecting off large white surfaces, to smooth lighting of the face. Simultaneous audio recordings were obtained using a Sennheiser ME64 microphone, placed at close proximity to the mouth (about 35 cm), connected to a PC via a MobilePre USB M-Audio pre-amplifier. During recordings, the talkers were instructed to look straight ahead with a neutral expression, say the numbers without using inflection or diphthongs and close their lips between utterances. To ensure a steady pace, a soft beeping noise was used as a trigger every 4 s. Recording of the sequence of numbers was repeated twice for each talker.

The same set of 24 lists designed for the updating task was created for both the female and male talkers. To create the lists, the externally recorded audio was firstly synchronized to the video by aligning the externally acquired audio signal with the audio signal recorded with the video camera using a cross-correlation method in MATLAB. This technique can align two signals to an accuracy within 0.02 ms. Subsequently, the audio signal of each number was normalized in level to the same nominal value after removing gaps in the speech. A MATLAB program was then used to cut the long clips into short clips that were joined together according to the specified list sequences. For each number, the better of the two takes was used. The joined audio/video segments were cross-faded to ensure a smooth transition in both audio and video. In the final lists, the spoken numbers occurred roughly every 2.5 s.

Finally, the audio was equalized per list to match the one-third octave levels of the ILTASS by Byrne et al. (1994).

Two kinds of background noise were used. One was an eight-talker babble noise from the National Acoustic Laboratories' CDs of Speech and Noise for Hearing Aid Evaluation (Keidser et al., 2002). This noise had low amplitude modulation and was filtered to match the ILTASS. The other noise was a simulated reverberant cafeteria scene (for a detailed description of the scene, see Best et al., 2015). In brief, the noise was simulated such that the listener is positioned amongst the seating arrangements of a cafeteria with the target talker having a virtual position in the room in front of the listener. The background consists of seven conversations between pairs of talkers seated at the surrounding tables and facing each other, resulting in 14 masker talkers distributed around the listener at different horizontal directions, distances and facing angles. Room impulse responses generated in ODEON (Rindel, 2000) were converted to loudspeaker signals using a loudspeaker-based auralisation toolbox (Favrot and Buchholz, 2010). This noise was more amplitude modulated than the babble-noise, but not as modulated as single-talker speech. To maintain its natural acoustic characteristics, it was not filtered to match the target material. Consequently, when equalized to the same Leq, the cafeteria noise exposed the target at frequencies above 1.5 kHz, see **Figure 1**.

Setup

Speech and noise were presented spatially separated in the free field using a 16-loudspeaker array in the horizontal plane of the listener's ears. The loudspeakers, Genelec 8020C active (self-amplified), were organized in a circle with a radius of 1.2 m and were driven by two ADI-8 DS digital-to-analog converters and an RME Fireface UFX interface, connected to a desktop PC. Using custom-made software, each loudspeaker was equalized (from 100 to 16000 Hz) and level-calibrated at the center of the array. The audio target was always presented from 0° azimuth at a level corresponding to 62 dB SPL at the position of the participant's head. The video signal of the CSCT was shown on a 21.5 inch PC monitor mounted on an independent stand and appearing above the frontal loudspeaker. As the video was presented at a resolution of 1440 × 1080 to a monitor supporting a resolution of 1920 × 1080, a black bar occurred on either side of the video. Four uncorrelated samples of the babble-noise were presented from ±45° azimuth and ±135° azimuth, while the reverberant cafeteria-noise was played back from all 16 loudspeakers. Custom-made menu-driven software was used to mix and present target and noise at specified SNR values in a real-time fashion. While the long-term levels of both target and noise were controlled, the short-term SNRs were not to maintain a natural interaction between target and noise. That is, the audibility of individual numbers likely varied within and between participants. Across all presentations, the effect of this variation is presumed to be leveled out. For the hearing-impaired participants, amplification was applied to all stimuli following the NAL-RP prescription (Byrne et al., 1990), with gain tapered to 0 dB at frequencies above 6 kHz. The prescribed filters were applied in real-time to the combined target and noise stimuli.

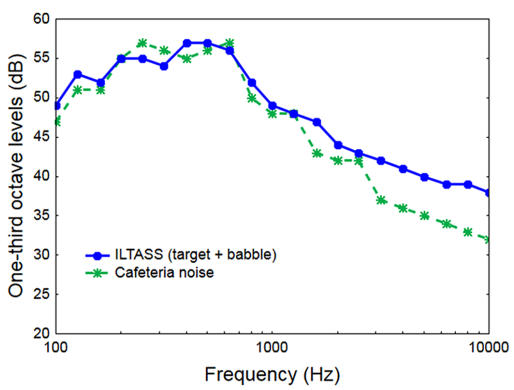


FIGURE 1 | The long-term spectra of the ILTASS (Byrne et al., 1994), that speech and babble-noise were filtered to match, and of the cafeteria-noise.

Cognitive Tests

The English version of the RST was adapted from Hällgren et al. (2001) as an independent test of working memory capacity. Sentences were presented on a screen in three parts and in spans of 3–6 sentences. Within each span, the inter-sentence interval was 3000 ms. After the end of every sentence; i.e., every third screen, the participants were asked to say 'yes' or 'no' to indicate whether that sentence was sensible or not. At the end of each span the participants were asked to recall either the first or last word of the sentences in that span. After a practice trial, 12 spans of sentences were presented, increasing from three series of three sentences to three series of six sentences.

The Letter Memory test (Morris and Jones, 1990) was used as an independent test of updating. An electronic version of the test was developed that presents 320 point size consonants on a screen, one by one, for a duration of 1 s each. Participants were presented with sequences of 5, 7, 9, or 11 consonants, and asked at the end of each sequence to recall the last four consonants. After two practice trials, three trials of each sequence length were presented in randomized order.

Protocol

Each participant attended one appointment of about 2 h. First, the purpose of the study and the tasks were explained, and a consent form was signed. Otoscopy was performed, followed by threshold measurements. The participants then completed the RST and the Letter Memory test. Both tests were scored manually, with the final scores comprising the percentage of correctly recalled words and letters, respectively, irrespective of order. This part of the appointment took place in a regular sound-treated test booth.

The remaining part of the appointment took place in a variable acoustic room, adjusted to a reverberation time of T_{60} = 0.3 s. Participants were seated in the center of the loudspeaker array. First they completed an adaptive speech-in-noise test to determine the individual SNR for testing CSC in noise. Using the automated, adaptive procedure described in Keidser et al. (2013), sensible high context sentences (filtered

to match the ILTASS) were presented in the eight-talker babble noise described above to obtain the SNR that resulted in 80% speech recognition. During the procedure the target speech was kept constant at 62 dB SPL while the level of noise was varied adaptively, starting at 0 dB SNR, based on the number of correctly recognized morphemes. Based on pilot data obtained on six normal-hearing listeners, the SNR was increased by 1 dB to reach the SNR that would result in ∼90% speech recognition when listening in babble-noise. This SNR was subsequently used in the CSCT with both the babble and cafeteria noises.

Finally, the CSCT was administered in a 2 (talker gender) × 3 (background noise, incl. quiet) × 2 (modality) design using two lists for each test condition. Test conditions were randomized in a balanced order across participants with lists further balanced across test conditions. After each list, participants had to recall either the two highest or the two lowest numbers in the list as instructed before each list. Because participants did not have to distinguish between talker gender while doing the updating task, a high memory load as introduced by Mishra et al. (2013a) was used; i.e., participants also had to remember the first number, as the task was otherwise considered too easy in the quiet condition for the younger normal-hearing listeners. The first number was not counted in the final score. During testing, participants verbalized their responses to the experimenter at the end of each list. Participants were instructed to look at the monitor during the audio-visual presentations, and this was reinforced by the experimenter who could observe the participants during testing. In the audio-only mode the video was switched off, meaning that the audio signal was the same in the two modalities.

Results and Discussion

Reading Span and Updating Tests

Table 1 lists the average performance data obtained by the two population groups on the reading span and updating tests. On both measures, the younger normal-hearing listeners outperformed the older hearing-impaired listeners. The differences in performance were significant according to a Mann–Whitney *U*-test ($p = 0.0005$ for the RST, and $p = 0.03$ for the updating test).

TABLE 1 | Mean and SE values for RST and updating test for each population group.

	Young normal-hearing		Older hearing-impaired	
Parameter	**Mean**	**SE**	**Mean**	**SE**
RST (%)	49.4	3.02	32.0	2.22
Updating (%)	84.5	2.06	76.2	3.07

Test Signal-to-Noise Ratios

Individually selected SNRs were obtained for testing CSC in noise. On average, the older hearing-impaired listeners needed higher SNRs (−1.0 dB; SE = 0.6 dB) than the younger normal-hearing listeners (−4.5 dB; SE = 0.4 dB). The difference in mean was significant according to a Mann–Whitney *U*-test ($p = 0.0001$).

Cognitive Spare Capacity

Figure 2 shows the average CSC score obtained by the younger and older listeners in each test condition. The arcsine transformed CSC scores were used as observations in a repeated measures ANOVA, using talker gender, noise, and modality as repeated measures and population group as grouping variable. This analysis revealed significant main effects of population group [$F(1,19) = 11.5$; $p = 0.003$], talker gender [$F(1,19) = 11.6$; $p = 0.003$], and noise [$F(2,38) = 6.5$; $p = 0.004$]. Specifically, the younger normal-hearing listeners showed more CSC than the older listeners across conditions, while CSC was reduced for the male talker (relative to the female talker) and by the presence of babble-noise (relative to quiet or cafeteria-noise). Modality did not show significance [$F(1,19) = 0.6$; $p = 0.46$], and none of the interactions were significant (p-levels varied from 0.08 for the three-way interaction of noise × modality × population group to 0.95 for the four-way interaction). Overall the English CSCT was sensitive to factors that could be expected to influence cognitive listening effort, although it differs from the Swedish

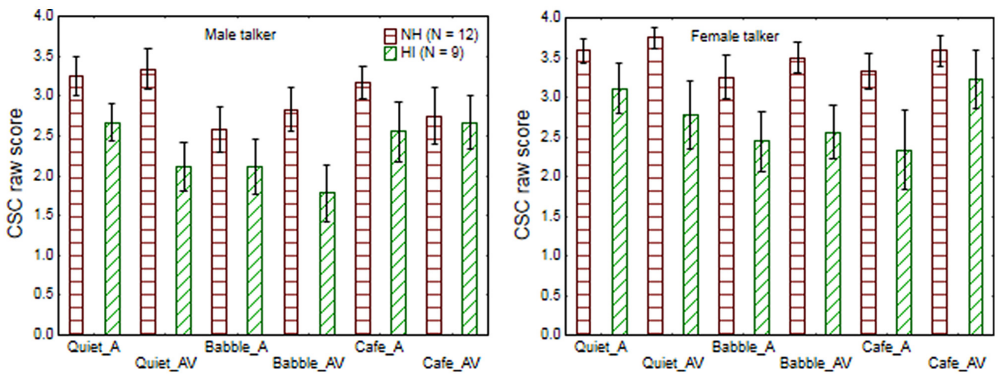


FIGURE 2 | The average CSC scores obtained by younger normal-hearing and older hearing-impaired participants when listening to a male talker (left graph) and female talker (right graph) in quiet, babble-noise (Babble), and in cafeteria-noise (Cafe) with audio-only (A) or audio-visual (AV) cues.

CSCT by not showing sensitivity to presentation modality, and no significant interaction between noise, modality, and population group.

The English version of the CSCT differed from the Swedish version by having more updating events as a result of presenting all numbers by a single talker instead of switching between two talkers. Targets were further presented in the free field instead of under headphones. **Table 2** shows the differences in average scores obtained with the English and Swedish versions of CSCT for comparable test conditions. As there were no significant interactions with talker gender, the CSC scores obtained for the English test were averaged across talker gender, while the CSC scores obtained for the Swedish test were eyeballed off the graphs in Mishra et al. (2013b, 2014). Our results obtained in the audio-only mode compared well with the results on the Swedish version of the CSCT, suggesting that the modifications introduced to the actual test had negligible effects on CSC.

On the independent visual tests, the older hearing-impaired listeners showed significantly reduced updating skill and working memory capacity compared to the younger normal-hearing listeners. These findings are in agreement with MacPherson et al. (2002) who found that age has a negative association with performance on tests of executive function and working memory. The older hearing-impaired listeners also showed significantly reduced CSC compared to the younger normal-hearing listeners, which agrees with Mishra et al. (2014). The two groups differed in hearing loss as well as age. Hearing loss, even when aided, would impact on speech understanding because of distortions such as temporal processing (Fitzgibbons and Gordon-Salant, 1996; Gordon-Salant and Fitzgibbons, 2001). However, differences in the amount of speech understood (caused by differences in speech understanding abilities due to hearing loss as well as cognitive ability) were removed by using individually selected SNRs. Therefore, the finding suggests that aging effects observed in executive and working memory processing extend to CSC, or mental effort. This agrees with Gosselin and Gagné (2011) who found that older adults generally expended more listening effort than young adults when listening in noise under equated performance conditions.

Relative to the female talker, our participants, on average, showed reduced CSC when listening to the male talker. When comparing the two talker materials, the female talker was notably more articulate than the male talker. Thus the significant gender effect likely occurred because clear production of speech, rather than the female voice *per se*, freed up cognitive resources in the listeners. This is in agreement with observations of Payton et al. (1994) and Ferguson (2004, 2012) who found that both normal-hearing and hearing-impaired listeners performed better on nonsense sentences and vowel identification, respectively, when listening to a speaking style that was deliberately made clear relative to a conversational version. Further research with a range of male and female talkers is needed to fully explore the effect of talker gender on cognitive listening effort in older hearing-impaired listeners.

On average, our listeners showed a significant reduction in CSC when listening in the babble-noise relative to listening in quiet, which is in line with findings for a stationary noise by Mishra et al. (2013b, 2014). While the hearing-impaired listeners in Mishra et al. (2014) also showed a reduction in CSC relative to quiet when listening in a highly modulated speech-like background noise, the normal-hearing listeners did not (Mishra et al., 2013b). Mishra et al. (2013b) have suggested that the younger listeners could take advantage of a selective attention mechanism that comes into play when speech is presented against a speech-like noise (Zion Golumbic et al., 2013) to track the target speech dynamically in the brain. In the stationary noise, it was argued, the absence of modulations reduced the ability to track the speech. For the older listeners, their less efficient cognitive functions made it more difficult to separate the target speech from the non-target speech, whether the noise was modulated or not. An alternative way to view this is that speech understanding for the two groups was equated only in the unmodulated noise. As is well-known, hearing-impaired listeners are less able to take advantage of gaps in a masker (Festen and Plomp, 1990; Hygge et al., 1992; Peters et al., 1998), so in the modulated noise, the hearing-impaired listeners would have had to apply more cognitive resources than the normal-hearing listeners just to understand the speech. Consequently, the normal-hearing listeners were less likely to have had their cognitive capacity depleted by the modulated noise than was the case for the hearing-impaired listeners. Overall, findings on the two versions of CSCT suggest that both normal-hearing and hearing-impaired listeners expend executive resources on hearing out the target from a noise that has a similar spectrum and thus exerts a uniform masking effect across all speech components. In our study, neither population group showed significantly reduced CSC when listening in cafeteria-noise relative to quiet. The individually selected test SNRs were obtained in babble-noise, and it is possible that because the cafeteria-noise was more speech-like than the babble-noise, at the same SNR, spatial separation would in this case have an effect. This notion is supported by several studies that have demonstrated that when target and maskers are spatially separated, it is relatively easier to extract speech from the less than the more distinguishable masker (Noble and Perrett, 2002; Arbogast et al., 2005). In addition, it is possible that better SNRs at high frequencies available in our cafeteria-noise made speech easier to access (Moore et al., 2010). Combined, these two factors may have made it easier for both population groups to identify and track the target speech, and hence

TABLE 2 | The difference in CSC scores obtained for the English and Swedish samples (English – Swedish) on comparable test conditions with an updating task presented under high memory load.

	Audio-only mode	Audio-visual mode
Normal-hearing		
quiet	−0.01	0.24
Noise with no or low modulation	0.06	0.13
Noise with high modulation	0.03	−0.21
Hearing-impaired		
quiet	0.09	−0.87
Noise with no or low modulation	0.29	0.09
Noise with high modulation	−0.09	0.39

reduce the cognitive resources needed for understanding, especially as our hearing-impaired listeners had very mild hearing loss.

The main discrepancy between the Swedish and English version of the CSCT is that the Swedish version was sensitive to presentation mode while the English version was not. With the Swedish version, older adults generally showed more CSC in the audio-visual mode relative to the audio-only mode (Mishra et al., 2014), whereas younger adults showed this pattern in noise but the opposite pattern when listening in quiet (Mishra et al., 2013a,b). The authors argued that under more demanding listening situations, the addition of visual cues counteracted the disruptive effect of noise and/or poorer hearing. This argument is supported by Frtusova et al. (2013) who found that visual cues facilitate working memory in more demanding situations for both younger and older adults, and Fraser et al. (2010) who saw a reduction in listening effort when introducing visual cues in a dual-task paradigm involving listening to speech in noise. For the younger cohort, the authors speculated that while listening in quiet, the auditory processing task was implicit, meaning that the visual input became a low priority stimulus and hence a distractor (Lavie, 2005), such that audio-visual integration required in the audio-visual mode added demand to the executive processing capacity. No effect of modality was observed in this study, which could suggest that our test conditions were not as cognitively demanding as those used by Mishra and colleagues although the data obtained in the audio-only mode in **Table 2** seem to refute this theory. Another possible reason for the lack of a visual effect in our study is poor attention to the video signal (Tiippana et al., 2004). Although the participants were all looking directly at the screen during testing, the room in which testing was conducted presented a lot of distracting visual information, including colorful wall panels, and the array of loudspeakers and other test equipment. Lavie (2005) has demonstrated that even when people have been specifically instructed to focus attention on a visual task, they are easily distracted while the perceptual load in the visual modality is low. Other data on the association between audio-visual integration and executive function are divided (Prabhakaran et al., 2000; Allen et al., 2006), hence, the visual effect on CSC needs a more systematic investigation.

The Association between CSC and Other Cognitive Measures

Regression analyses were performed to investigate the association between the factor-wise CSC scores (i.e., scores averaged across various experimental conditions) obtained on all participants and the other two cognitive measures, when either controlling for 4FA HL or age. Separate regression analyses were performed using each of the reading span and updating measures as independent variable. The results are summarized in **Table 3**. In all cases, the regression coefficient was positive, sometimes significantly so; suggesting that more CSC was associated with better cognitive function. The results were little affected whether age or hearing loss was used as the co-variate. In agreement with Mishra et al. (2013a,b, 2014), the CSCT was more strongly related to the updating test than to the RST. Overall, the more consistent association with the independent updating test and inconsistent association with the RST suggest that the CSCT measures something more similar to the combination of attributes used in the updating task than those used in the RST. However, for none of the individual CSC scores is the association between CSC and updating skill significantly greater than the association between CSC and reading span measures. We further note that moderate, but significant, correlations have been found between measures of memory updating and complex working memory spans (e.g., Lehto, 1996).

Experiment II

The aims of Experiment II were to examine, in normal-hearing listeners, if CSCT or RST measures would better predict comprehension of dynamic conversations, and if CSC is reduced when increasing the dynamics of the listening situation. Speech performance was measured using a new speech comprehension test that delivers monologs and conversations between 2 and 3 spatially separated talkers. Participants listened to the speech and

TABLE 3 | The standardized regression coefficients (ß) and their SE values related to the extent to which CSC scores are predicted by performance on the RST or updating test when controlling for degree of hearing loss (4FA HL) or age.

	RST (%)				Updating test (%)			
	4FA HL (dB HL)		Age (year)		4FA HL (dB HL)		Age (year)	
Parameter	ß	SE of ß	ß	SE of ß	ß	SE of ß	ß	SE of ß
CSCT overall	0.50	0.23	0.38	0.23	0.59**	0.18	0.54**	0.18
Male	0.59*	0.25	0.54*	0.23	0.60**	0.18	0.60**	0.19
Female	0.30	0.29	0.15	0.24	0.45*	0.21	0.38	0.20
Quiet	0.17	0.27	0.13	0.23	0.57**	0.17	0.51**	0.17
Cafeteria	0.70*	0.26	0.52	0.25	0.56*	0.20	0.55*	0.21
Babble	0.43	0.27	0.34	0.23	0.41*	0.20	0.36	0.20
A-only	0.26	0.30	0.11	0.25	0.54*	0.20	0.48*	0.20
AV	0.65*	0.23	0.58*	0.20	0.55**	0.17	0.53*	0.18

One asterisk indicates a significance level <0.05, and two asterisks a significance level <0.01.

answered questions about the information while continuing to listen. To parallel the dynamic speech comprehension test, the CSCT stimuli were presented either all from a single loudspeaker position, or randomly from two or three loudspeaker positions. Both the CSCT and the dynamic speech comprehension test were implemented under realistic acoustic conditions in a cafeteria background.

Considering the mental processes involved in performing the RST (reading words, deriving meaning from the words, forming and delivering a response, storing items, and recalling items), the CSCT (segregating target speech from noise, recognizing the words, making decision about what to store, storing items, deleting items, and recalling items), and the speech comprehension test (segregating target speech from noise, recognizing, the words, deriving meaning from the words, storing items, recalling items and forming and delivering a response,), it would seem that the speech comprehension test shares processes with both the RST and the CSCT, and that only a couple of operations are common to all three tests. Based on a comparison of the mental processes the pairs of tests have in common, it could be expected that speech comprehension performance would be more correlated with performance on the RST if individual differences in the ability to process words to derive meaning and form a response are more important in causing individual differences in speech comprehension than individual differences in identifying which speech stream is the target, segregating it, and recognizing the words. With our group of normal-hearing listeners we expected the former to be the case and hence we predicted performances on our comprehension test to be associated more strongly with RST than with CSCT measures. We further expected that increasing the dynamic aspects of speech by changing voice and location of talkers more frequently would add processing demands in working memory, and in the executive function specifically, so that the listeners would require better SNRs to perform as well in the conversations as in the monologs (Kirk et al., 1997; Best et al., 2008), and that between listening conditions, variations in the CSC would be correlated with variations in speech comprehension.

Methodology

Participants

The participants were primarily university students and included 16 females and 11 males. All had normal hearing, showing an average 4FA HL of 2.9 dB HL (SE = 0.6 dB). The age of the participants ranged from 18 to 40 years, with an average of 26.2 years. Participants were paid a small gratuity for their inconvenience.

Dynamic Speech Comprehension Test

The dynamic speech comprehension test consists of 2–4 min informative passages on everyday topics that are delivered as monologs or conversations between two or three talkers. The passages are taken from the listening comprehension component of the International English Language Testing System, for which transcripts and associated comprehension questions are publicly available in books of past examination papers (Jakeman and McDowell, 1995). The recorded presentations are spoken by voice-actors who were instructed to read the monologs and play out the conversations in a natural way, including variations in speed, pauses, disfluencies, interjections etc. Each passage is associated with 10 questions that are answered "on the go" (brief written responses) while listening.

Setup

Testing took place in an anechoic chamber fitted with 41 equalized Tannoy V8 loudspeakers distributed in a three-dimensional array of radius 1.8 m. In the array, 16 loudspeakers were equally spaced at 0° elevation, eight at ±30° elevation, four at ±60° elevation, and one loudspeaker was positioned directly above the center of the array. Stimuli were played back via a PC equipped with an RME MADI soundcard connected to two RME M-32 D/A converters and 11 Yamaha XM4180 four-channel amplifiers.

Testing was done in a simulated cafeteria scene similar to that used in Experiment I. The background noise was simulated using ODEON software (Rindel, 2000) in the same way as described for the cafeteria noise in Experiment I, but using different room characteristics, and the entire 41 loudspeaker array. As previously, the background of the cafeteria noise consisted of seven conversations between pairs of talkers seated at tables and facing each other, resulting in 14 masker talkers distributed around the listener at different horizontal directions, distances and facing angles. The listener was situated by a table slightly off center in the room, facing three talkers positioned 1 m away at −67.5, 0, and +67.5° azimuth. During testing, monologs were presented from either of these three loudspeaker locations. For the two-talker condition, conversations took place between talkers situated at −67.5 and 0°, at 0 and +67.5°, or at −67.5 and +67.5° azimuths. The three-talker conversations all involved the talkers at each of the three loudspeaker locations. While speech was presented from each of these loudspeakers, an LED light placed on top of the loudspeaker was illuminated to give the listener a simple visual cue to indicate which source was active, as would be indicated by facial animation and body language in a real conversation.

Protocol

Each participant attended three appointments of about 2 h. During the first appointment, the purpose of the study and the tasks were explained, and a consent form was signed. Otoscopy was performed, followed by threshold and reading span measurements. The implementation of the RST was the same as used in Experiment I. The dynamic speech comprehension test was completed over the three appointments, and the CSCT was administered at either the second or third appointment.

For the dynamic speech comprehension test, the target speech was fixed at 65 dB SPL and all participants were tested in each talker condition at three SNRs (−6, −8, and −10 dB), using five passages (i.e., 50 scoring units) for each SNR. The participant was seated in the anechoic chamber such that the head was in the center of the loudspeaker array, facing the frontal loudspeaker. Note that participants were allowed to move their head during testing to face the active source. Responses were provided in written form using paper and pencil and scored manually post-testing. The different passages were balanced across test conditions, and talker conditions and SNRs were presented in a randomized

order across participants. The source position of the talkers also varied randomly across and within passages.

The CSCT was presented in a similar fashion to the dynamic speech test at a −6 dB SNR. Three lists were administered for each talker condition and the combined score obtained. To parallel the one-talker condition, one list was presented from each of the three talker locations (−67.5, 0, and +67.5° azimuths). To parallel the two-talker condition, numbers were for one list randomly presented from −67.5° and 0° azimuths, for another list randomly presented from 0 and +67.5° azimuths, and for the final list randomly presented from −67.5 and +67.5° azimuths. To parallel the three-talker condition, numbers for each of the three lists were randomly presented from the three loudspeaker locations. To reduce the chance of reaching ceiling effects, a high memory load was implemented by asking the participants to also recall the first number in each list, although the number was not counted in the final score. Before CSC testing, one list was presented in −6 dB SNR, with numbers coming randomly from two loudspeaker locations, and participants were asked to repeat back the numbers heard. One missed number was allowed; otherwise the SNR was increased to ensure that the participants were able to hear the numbers in the noise. No participants needed the SNR changed. Nine lists from a pool of 12 were randomly selected for each participant and randomly presented across talker condition and locations.

Results and Discussion

Speech Comprehension

For each participant a logistic function was fitted to the three data points measured with the comprehension test for each talker condition, and the SNR for 70% correct answers was extracted (SRT_{70}). For three participants, the data obtained for one talker condition (single-talker or three-talker) were not well-behaved as a function of SNR, and thus sensible logistic functions could not be fit. From the remaining 24 participants, the average differences in SRT_{70} between the 1- and 2-talker, and between the 2- and 3-talker conditions, were obtained. These differences were applied as appropriate to the two-talker SRT_{70} values measured for the three participants with missing data points to obtain extrapolated replacement values. According to a repeated measures ANOVA the difference in SRT_{70} between talker conditions was significant [$F(2,52) = 3.92$; $p = 0.03$], **Figure 3**. A Tukey HSD *post hoc* analysis revealed that the listeners required significantly higher SNRs to reach 70% correct scores on the monologs than on the dialogs. We note that the ranking of conditions in terms of SRTs corresponds to the complexity of the language of the passages, as measured with the Flesch–Kincaid Grade level (Kincaid et al., 1975; 9.7, 3.5, and 6.1 for the one, two, and three-talker passages, respectively). This suggests that speech comprehension may be more affected by complexity of the spoken language, in terms of length and number of words used, than by the dynamic variation in talker location.

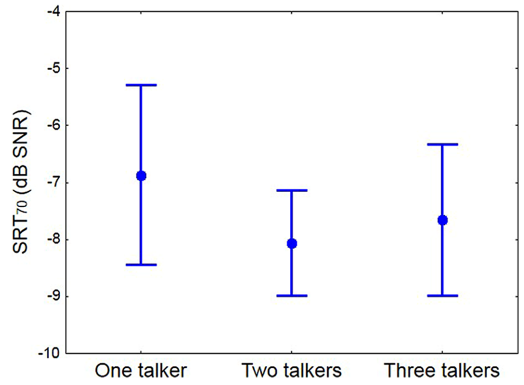


FIGURE 3 | The mean SRT_{70} for each talker condition. Whiskers show 95% confidence interval.

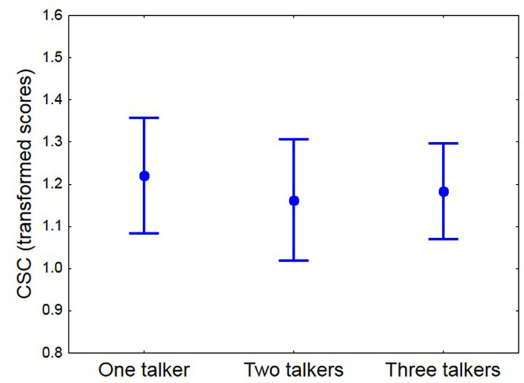


FIGURE 4 | The mean transformed CSC score for each simulated talker condition (maximum = 1.57). Whiskers show 95% confidence interval.

The Sensitivity of CSC to Increased Dynamic Variation

To investigate if CSC was affected by increasing the number of talkers in the listening situation, the combined scores across three CSC lists were obtained for each participant and simulated talker condition. Based on arcsine transformed scores, participants, on average, showed slightly reduced CSC for the simulated two-talker condition relative to the simulated one- and three-talker conditions, **Figure 4**. According to a repeated measures ANOVA this pattern was not significant [$F(2,52) = 0.27$; $p = 0.76$], suggesting that, at least for younger normal-hearing listeners, increasing the complexity of the listening condition, by increasing the number of target locations, did not reduce CSC. It is worth noting, that the lowest average CSC of 1.1 transformed scores was obtained for the two-talker condition in which the target locations were most separated (by 67.5°).

Predicting Inter-Participant Variation in Speech Comprehension

Across participants, reading span scores varied from 28 to 70% with a mean of 45.5%. This result is not unlike findings by Zekveld et al. (2011), who reported a mean reading span score of 48.3%, ranging from 30 to 74%, on a slightly younger normal-hearing sample. **Table 4** lists the correlation coefficients for the associations between reading span scores and transformed CSC scores obtained for each talker condition (first column). Reading

TABLE 4 | The correlation coefficients and their 95% confidence intervals (shown in brackets) for associations of interest between RST, CSCT, and SRT_{70} measures.

Parameter	RST	SRT_{70} (1-talker)	SRT_{70} (2-talker)	SRT_{70} (3-talker)	SRT_{70} (collapsed)
RST		−0.51** [−0.15,−0.75]	−0.52** [−0.17,−0.76]	−0.57** [−0.24,−0.78]	−0.57** [−0.24,−0.78]
CSCT (1-talker)	0.04 [0.42,−0.35]	−0.05 [0.34,−0.43]			
CSCT (2-talker)	0.4* [0.69,0.04]		−0.29 [0.11,−0.61]		
CSCT (3-talker)	−0.08 [0.32,−0.45]			0.09 [0.46,−0.31]	
CSCT (collapsed)	0.21 [0.55,−0.19]				−0.21 [0.19,−0.55]

One asterisk indicates a significance level <0.05, and two asterisks a significance level <0.01.

span scores were positively and significantly associated with the transformed CSC scores obtained for the simulated two-talker condition ($p = 0.03$), but not for the simulated one- and three-talker conditions ($p = 0.83$ and $p = 0.69$, respectively). The fact that CSC scores are not consistently correlated with reading span measures across all three conditions may suggest again that the two tests do not generally capture the same cognitive constructs, although none of the correlation coefficients were significantly different from each other.

To determine whether CSCT or RST best predicted inter-participant variation in speech comprehension, correlation coefficients for the association between reading span scores and performance on the speech comprehension test in each talker condition (first row), and for each talker condition the association between transformed CSC scores and performance on the speech comprehension test were obtained, see **Table 4**. For all three talker conditions, data suggest that good performance on the dynamic speech comprehension test requires good working memory capacity ($p < 0.01$ for all three talker conditions), but is not significantly associated with cognitive listening effort as measured with the CSCT ($p = 0.82$, $p = 0.15$, and $p = 0.67$ for the one-, two-, and three-talker condition, respectively). As associations between measures were consistent across talker conditions, data for the CSCT and speech comprehension measures were further collapsed across talker conditions to do an overall three-way correlation analysis. As can be seen in **Table 4**, the association between RST and the collapsed SRT_{70} is highly significant ($p = 0.002$), while the association between the collapsed CSC and SRT_{70} is not ($p = 0.30$). The difference between the correlation coefficients obtained for the two associations is, however, not significant ($p = 0.13$), meaning that no strong conclusion can be made about the relative strengths of the associations. Looking at the three-way correlation matrix, where the association between the collapsed CSC scores and RST is also non-significant ($p = 0.31$), it is evident, however, that the strongest similarity is found between the SRT_{70} and RST measures.

Overall Discussion

Two experiments were presented in this paper. In the first experiment we evaluated an English version of the CSCT introduced by Mishra et al. (2013a) that focuses on measuring an individual's CSC for updating processing after processing of auditory stimuli has taken place. In the second experiment we investigated if this measure of CSC or a measure of working memory capacity, using the RST, better predicted variation in speech comprehension, and if CSC was reduced when increasing the number of talkers in the listening situation.

In agreement with Mishra et al. (2013a,b, 2014) we found in both experiments indications that the CSCT measures a construct different from the RST. This was expected as the two test paradigms do differ in some of the mental processes that are required to perform the specific tasks of the tests. The evidence was, however, not strong. Specifically, we note that with an administration of two lists per test condition, 74% of variance in CSC scores obtained in Experiment I was due to intra-participant measurement error variance, which would have reduced the reported regression coefficients. Further, there is some concern to what extent participants actively engage in updating when the task is to recall the last items in a list of an unknown number of items, as is the case in the independent updating task employed in Experiment I, or whether they simply wait until the end of the list before attempting to recall the most recent items (Palladino and Jarrold, 2008). Consequently, the correlation analyses presented in this study and in Mishra et al. (2013b, 2014) on the associations between the RST and the CSCT scores and between the independent updating task and the CSCT scores should be interpreted with caution. Overall, it would be desirable in the future to establish the psychometric properties of the CSCT, including determining the ideal number of lists for reliable measures of CSC, and to more systematically explore the relationship between CSCT, RST, and other tests of executive processing and working memory capacity.

Evaluated in a more natural listening environment than that used by Mishra et al. (2013a,b, 2014), we confirmed in Experiment I that the CSCT has merit as a concept for measuring the cognitive effort associated with listening to speech that has been degraded by some form of distortion. Specifically, we found that the CSCT was sensitive to population group and a masker with low modulation (relative to listening in quiet), and further to clarity of speech. On the other hand, we could not confirm in Experiment I that CSC is affected by a masker with high modulation in hearing-impaired listeners or by presentation modality in either population group. Methodological variations are suggested to account for the differences observed between the English and Swedish version of the CSCT. Specifically, spatial separation of target and masker, and exposure to high-frequency speech energy when listening in the highly modulated cafeteria-noise likely made it easier for both population groups to access and track target speech (Arbogast et al., 2005; Moore et al., 2010), and

hence in line with the ELU model made this test condition less taxing on cognitive effort. A low perceptual load in the visual modality and distracting visual information in the test environment were suggested to combine to have made participants prone to relax their attention to the video signal (Tiippana et al., 2004; Lavie, 2005), to reduce its potential effect on cognitive listening effort. It would be of interest to study these factors more closely in the future. It should also be noted that if our implementations indeed were closer to real-life listening, this study would suggest that cognitive listening effort may not be as easily modulated by the listening condition in real life as demonstrated in some laboratory tests.

As predicted on the basis of the mental processes involved in our speech comprehension test, and our participant sample having normal-hearing, we found in Experiment II that those with poorer working memory capacity required better SNRs to perform at a similar level on the comprehension test than those with greater capacity. The association between speech comprehension and working memory capacity was significant, while the association between speech comprehension and CSC was not, suggesting that individual differences in speech comprehension may be more related to individual abilities to process words to derive meaning and form a response than to the individual abilities to overcome the perceptual demand of the task. This finding ties in well with the established association between span tests, such as the RST that tap into the combined processing and storage capacity of working memory, and speech comprehension (Daneman and Merikle, 1996; Waters and Caplan, 2005), and further lends support to the ELU model. We speculate, however, that we may see an opposite trend in a hearing-impaired population; i.e., find a significant association between speech comprehension and CSC instead. This is because the individual abilities in this population to meet the perceptual demands of the CSCT may outweigh the variation in individual abilities to process written words to derive meaning and form a response.

The finding in Experiment II that increasing the dynamic variation in voice and location from 1 to 2 and three talkers did not systematically affect speech comprehension performance in young normal-hearing participants, when they listened in a reverberant cafeteria-like background, was somewhat surprising. We had expected that the participants would have required slightly better SNRs for comprehending speech when listening to more than one talker (Kirk et al., 1997; Best et al., 2008) as turn-taking becomes less predictable, increasing the challenge of identifying the current talker and monitoring and integrating what each talker said. That is, they needed to expend more cognitive resources when listening to the conversations. However, it is possible that the increased cognitive demand arising from applying attention to location was counteracted by advantages from having a greater number of discourse markers and more informative perspectives from multiple talkers in the multi-talker conversations (Fox Tree, 1999). A significantly higher SRT_{70} measured for monologs than for dialogs may be explained by more and longer words being presented in the monologs than in the two-person conversations. This finding is in line with other studies that have seen sentence complexity impacting on speech comprehension performances (Tun et al., 2010; Uslar et al., 2013). The theory is also supported by findings that longer words reduce memory spans of sequences of words (Mueller et al., 2003); i.e., demand more working memory processing. However, we saw no difference in the strengths of the associations between RST scores and speech comprehension across talker conditions (cf. **Table 4**).

Previous studies have shown that measures of cognitive effort can be more sensitive to subtle changes in the listening situation than measures of speech understanding (e.g., Sarampalis et al., 2009; Ng et al., 2013). Thus, we expected that the CSCT might be sensitive to dynamic variations in target location even where our comprehension task was not. However, we found in Experiment II that applying random dynamic variations to the speech targets of the CSCT did not generally lead to reduced CSC in our normal-hearing participants, although it is of interest that the average lowest CSC was observed for the condition when numbers were presented randomly from the two most distant locations. Despite using transformed CSC scores in our analysis, our result may be partly influenced by many listeners reaching ceiling on the CSCT across test conditions (35% of total scores). It is also possible that allowing listeners to naturally move their head to listen to the spatially separated targets reduced differences in CSC, especially when distances between target locations were less extreme. On the other hand, it appeared from spontaneous comments that at least for some participants the shifting location of the target did not interfere with the task of updating the heard input, and thus it is possible that dynamic changes in target location did not actually represent a change in difficulty. It is worth noting that in the CSCT the actual voice did not change with location as it did in the dynamic speech comprehension test.

Future studies in our laboratory will further investigate to what extent CSC is sensitive to increasing complexity in the environment, and will also examine the effect of age and hearing loss on associations between CSC and the listening environment.

Acknowledgments

The work presented in this paper was partly sponsored by a grant from the Hearing Industry Research Consortium (IRC) and by the Australian Government Department of Health. Virginia Best was also partially supported by NIH/NIDCD grant DC04545. The authors would like to thank their colleagues Chris Oreinos, Adam Westermann, and Jörg Buchholz for helping out with recording and editing speech and noise stimuli, writing applications to control playback of stimuli, and calibrating the test setups.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Allen, R. J., Baddeley, A. D., and Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *J. Exp. Psychol. Gen.* 135, 298–313. doi: 10.1037/0096-3445.135.2.298

Arbogast, T. L., Mason, C. R., and Kidd, G. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 117, 2169–2180. doi: 10.1121/1.1861598

Arehart, K. H., Souza, P., Baca, R., and Kates, J. M. (2013). Working memory, age, and hearing loss: susceptibility to hearing aid distortion. *Ear Hear.* 34, 251–260. doi: 10.1097/AUD.0b013e318271aa5e

Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Baddeley, A. D., and Hitch, G. (1974). Working memory. *Psychol. Learn. Motiv.* 8, 47–89. doi: 10.1016/S0079-7421(08)60452-1

Best, V., Keidser, G., Buchholz, J. M., and Freeston, K. (2015). An examination of speech reception thresholds measured in a simulated reverberant cafeteria environment. *Int. J. Audiol.* doi: 10.3109/14992027.2015.1028656 [Epub ahead of print].

Best, V., Ozmeral, E. J., Kopčo, N., and Shinn-Cunningham, B. G. (2008). Object continuity enhances selective auditory attention. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13173–13177. doi: 10.1073/pnas.0803718105

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., et al. (1994). An international comparison of long-term average speech spectra. *J. Acoust. Soc. Am.* 96, 2108–2120. doi: 10.1121/1.410152

Byrne, D., Parkinson, A., and Newall, P. (1990). Hearing aid gain and frequency response requirements for the severely/profoundly hearing-impaired. *Ear Hear.* 11, 40–49. doi: 10.1097/00003446-199002000-00009

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 1559–1569. doi: 10.1016/S0022-5371(80)90312-6

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546

Favrot, S., and Buchholz, J. M. (2010). LoRA – A loudspeaker-based room auralization system. *Acta Acust. United Acust.* 96, 364–376. doi: 10.3813/AAA.918285

Ferguson, S. H. (2004). Talker differences in clear and conversational speech: vowel intelligibility for normal-hearing listeners. *J. Acoust. Soc. Am.* 116, 2365–2373. doi: 10.1121/1.1788730

Ferguson, S. H. (2012). Talker differences in clear and conversational speech: vowel intelligibility for older adults with hearing loss. *J. Speech Lang. Hear. Res.* 55, 779–790. doi: 10.1044/1092-4388(2011/10-0342)

Festen, J. M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 88, 1725–1736. doi: 10.1121/1.400247

Fitzgibbons, P. J., and Gordon-Salant, S. (1996). Auditory temporal processing in elderly listeners. *J. Am. Acad. Audiol.* 7, 183–189.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Process.* 27, 35–53. doi: 10.1080/01638539909545049

Fraser, S., Gagne, J.-P., Alepins, M., and Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech cues. *J. Speech Lang. Hear. Res.* 53, 18–33. doi: 10.1044/1092-4388(2009/08-0140)

Frtusova, J. B., Winneke, A. H., and Phillips, N. A. (2013). ERP evidence that auditory –visual speech facilitates working memory in younger and older adults. *Psychol. Ageing* 28, 481–494. doi: 10.1037/a0031243

Gordon-Salant, S., and Fitzgibbons, P. J. (2001). Source of age-related recognition difficulty for time-compressed speech. *J. Speech Lang. Hear. Res.* 44, 709–719. doi: 10.1044/1092-4388(2001/056)

Gosselin, P. A., and Gagné, J.-P. (2011). Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *Int. J. Audiol.* 50, 786–792. doi: 10.3109/14992027.2011.599870

Hällgren, M., Larsby, B., Lyxell, B., and Arlinger, S. (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. *J. Am. Acad. Audiol.* 12, 357–370.

Helfer, K. S. (1995). "Auditory perception by older adults," in *Communication in Later Life*, eds R. A. Huntley and K. S. Helfer (Boston, MA: Butterworth – Heinemann), 41–84.

Helfer, K. S., and Freyman, R. L. (2004). The role of visual speech cues in reducing energetic and informational masking. *J. Acoust. Soc. Am.* 117, 842–849. doi: 10.1121/1.1836832

Hygge, S., Rönnberg, J., Larsby, B., and Arlinger, S. (1992). Normal-hearing and hearing-impaired subjects' ability to just follow conversation in competing speech, reversed speech, and noise backgrounds. *J. Speech Hear. Res.* 35, 208–215. doi: 10.1044/jshr.3501.208

Jakeman, V., and McDowell, C. (1995). *Cambridge English IELTS 1 With Answers: Authentic Examination Papers from Cambridge ESOL*. Cambridge: Cambridge University Press.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149. doi: 10.1037/0033-295X.99.1.122

Keidser, G., Ching, T., Dillon, H., Agung, K., Brew, C., Brewer, S., et al. (2002). The National Acoustic Laboratories' (NAL) CDs of speech and noise for hearing aid evaluation: normative data and potential applications. *Aust. N. Z. J. Audiol.* 24, 16–35. doi: 10.1375/audi.24.1.16.31112

Keidser, G., Dillon, H., Mejia, J., and Ngyuen, C. V. (2013). An algorithm that administer adaptive speech-in-noise testing to a specified reliability at any point on the psychometric function. *Int. J. Audiol.* 52, 795–800. doi: 10.3109/14992027.2013.817688

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., and Chissom, B. S. (1975). *Deviation of a New Readability Formula for Navy Enlisted Personnel*. Millington, TN: Navy Research Branch.

Kirk, K. I., Pisoni, D. B., and Miyamoto, R C. (1997). Effects of stimulus variability on speech perception in listeners with hearing impairment. *J. Speech Lang. Hear. Res.* 40, 1395–1405. doi: 10.1044/jslhr.4006.1395

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends Cogn. Sci.* 9, 75–82. doi: 10.1016/j.tics.2004.12.004

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Q. J. Exp. Psychol. A* 49, 29–50. doi: 10.1080/713755616

Luce, P. A., and Pisoni, D. A. (1998). Recognising spoken words: the neighbourhood activation model. *Ear Hear.* 19, 1–36. doi: 10.1097/00003446-199802000-00001

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42(Suppl. 1), S49–S58. doi: 10.3109/14992020309074624

Lyxell, B., Andersson, U., Borg, E., and Ohlsson, I. S. (2003). Working-memory capacity and phonological processing in deafened adults and individuals with a severe hearing impairment. *Int. J. Audiol.* 42(Suppl. 1), S86–S89. doi: 10.3109/14992020309074628

MacPherson, S. E., Phillips, L. H., and Della Sala, S. (2002). Age, executive function and social decision making: a dorsolateral prefrontal theory of cognitive aging. *Psychol. Aging* 17, 598–609. doi: 10.1037/0882-7974.17.4.598

Marslen-Wilson, W. D. (1990). "Activation, competition, and frequency in lexical access," in *Cognitive Models of Speech Processing: Psycholinguistics and Computational Perspectives*, ed. G. T. M. Altmann (Cambridge, MA: MIT Press).

Marslen-Wilson, W. D., and Tyler, L. K. (1980). The temporal structure of spoken language comprehension. *Cognition* 6, 1–71. doi: 10.1016/0010-0277(80)90015-3

McCabe, D. P., Roediger, H. L. III, McDaniel, M. A., Balota, D. A., and Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology* 24, 222–243. doi: 10.1037/a0017619

McClelland, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cogn. Psychol.* 23, 1–44. doi: 10.1016/0010-0285(91)90002-6

McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Psychol.* 18, 695–698. doi: 10.1016/0010-0285(86)90015-0

Meister, H., Schreitmüller, S., Grugel, L., Ortmann, M., Beutner, D., Walger, M., et al. (2013). Cognitive resources related to speech recognition with a competing talker in young and older listeners. *Neuroscience* 232, 74–82. doi: 10.1016/j.neuroscience.2012.12.006

Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., and Runder, M. (2013a). Visual information can hinder working memory processing of speech. *J. Speech Lang. Hear. Res.* 56, 1120–1132. doi: 10.1044/1092-4388(2012/12-0033)

Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., and Runder, M. (2013b). Seeing the talker's face supports executive processing of speech in steady state noise. *Front. Syst. Neurosci.* 7:96. doi: 10.3389/fnsys.2013.00096

Mishra, S., Stenfelt, S., Lunner, T., Rönnberg, J., and Runder, M. (2014). Cognitive spare capacity in older adults with hearing loss. *Front. Aging Neurosci.* 6:96. doi: 10.3389/fnagi.2014.00096

Miyake, A., Friedman, N. P., Emerson, M. J., Witziki, A. H., Howerter, A., and Wager, T. (2000). The unity and diversity of executive functions and their contribution to complex frontal lobe tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (2010). Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task. *J. Acoust. Soc. Am.* 128, 360–371. doi: 10.1121/1.3436533

Morris, N., and Jones, D. M. (1990). Memory updating in working memory: the role of the central executive. *Br. J. Psychol.* 81, 111–121. doi: 10.1111/j.2044-8295.1990.tb02349.x

Mueller, S. T., Seymour, T. L., Kieras, D. E., and Meyer, D. E. (2003). Theoretical implications of articulartory duration, phonological similarity, and phonological complexity kin verbal working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1353–1358. doi: 10.1037/0278-7393.29.6.1353

Ng, E. H., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.* 52, 433–441. doi: 10.3109/14992027.2013.776181

Noble, W., and Perrett, S. (2002). Hearing speech against spatially separate competing speech versus competing noise. *Percept. Psychophys.* 64, 1325–1336. doi: 10.3758/BF03194775

Palladino, P., and Jarrold, C. (2008). Do updating tasks involve updating? Evidence from comparisons with immediate serial recall. *Q. J. Exp. Psychol.* 61, 392–399. doi: 10.1080/17470210701664989

Payton, K. L., Uchanski, R. M., and Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *J. Acoust. Soc. Am.* 95, 1581–1592. doi: 10.1121/1.408545

Peters, R. W., Moore, B. C. J., and Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. Acoust. Soc. Am.* 103, 577–587. doi: 10.1121/1.421128

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Pinquart, M., and Sörensen, S. (2000). Influences of socioeconomic status, social network, and competence on subjective well-being in later life: a meta-analysis. *Psychol. Aging* 15, 187–224. doi: 10.1037/0882-7974.15.2.187

Prabhakaran, V., Narayanan, K., Zhao, Z., and Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nat. Neurosci.* 3, 85–90. doi: 10.1038/71156

Repovš, G., and Baddeley, A. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience* 139, 5–21. doi: 10.1016/j.neuroscience.2005.12.061

Rindel, J. H. (2000). The use of computer modeling in room acoustics. *J. Vibroengineering* 3, 41–72.

Rönnberg, J., Arlinger, S., Lyxell, B., and Kinnefors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive function. *J. Speech Hear. Res.* 32, 725–735. doi: 10.1044/jshr.3204.725

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S99–S105. doi: 10.1080/14992020802301167

Rudner, M., Ronnberg, J., and Lunner, T. (2011a). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4

Rudner, M., Ng, E. H., Rönnberg, N., Mishra, S., Rönnberg, J., Lunner, T., et al. (2011b). Cognitive spare capacity as measure of listening effort. *J. Hear. Sci.* 1, 1–3.

Sarampalis, A., Kalluri, S., Edwards, B., and Hafter, E. (2009). Objective measures of listening effort: effects of background noise and noise reduction. *J. Speech Lang. Hear. Res.* 52, 1230–1240. doi: 10.1044/1092-4388(2009/08-0111)

Schneider, B. A., Li, L., and Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *J. Am. Acad. Audiol.* 18, 559–572. doi: 10.3766/jaaa.18.7.4

Schneider, B. A., Pichora-Fuller, M. K., and Daneman, M. (2010). "Effects of senescent changes in audition and cognition on spoken language comprehension," in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (New York, NY: Springer), 167–210.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., and Lewis, D. E. (2001). Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *J. Am. Acad. Audiol.* 110, 2183–2190. doi: 10.1121/1.1400757

Tiippana, K., Andersen, T. S., and Sams, M. (2004). Visual attention modulates audiovisual speech perception. *Eur. J. Cogn. Psychol.* 16, 457–472. doi: 10.1080/09541440340000268

Tun, P. A., Benichov, J., and Wingfield, A. (2010). Response latencies in auditory sentence comprehension: effects of linguistic versus perceptual challenge. *Psychol. Aging* 25, 730–735. doi: 10.1037/a0019300

Uslar, V. N., Carroll, R., Hanke, M., Hamann, C., Ruigendijk, E., Brand, T., et al. (2013). Development and evaluation of a linguistically and audiologically controlled sentence intelligibility test. *J. Am. Acad. Audiol.* 134, 3039–3056. doi: 10.1121/1.4818760

Waters, G., and Caplan, D. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory* 13, 403–413. doi: 10.1080/09658210344000459

Wingfield, A., and Tun, P. A. (2007). Cognitive supports and cognitive constraints on comprehension of spoken language. *J. Am. Acad. Audiol.* 18, 548–558. doi: 10.3766/jaaa.18.7.3

Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., van Beek, J. H. M., and Ronnberg, J. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear Hear.* 32, E16–E25. doi: 10.1097/AUD.0b013e318228036a

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "Cocktail Party." *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Keidser, Best, Freeston and Boyce. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 22 December 2014
doi: 10.3389/fpsyg.2014.01490

Memory performance on the Auditory Inference Span Test is independent of background noise type for young adults with normal hearing at high speech intelligibility

Niklas Rönnberg [1,2] ****, Mary Rudner*** [2,3] ***, Thomas Lunner*** [1,2,3,4] ***and Stefan Stenfelt*** [1,2]

[1] *Technical Audiology, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden*
[2] *Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden*
[3] *Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden*
[4] *Oticon Research Centre Eriksholm, Snekkersten, Denmark*

Edited by:
Claude Alain, Rotman Research Institute, Canada

Reviewed by:
Elvira Brattico, University of Helsinki, Finland
Kristina C. Backer, University of Washington, USA

****Correspondence:***
Niklas Rönnberg, Technical Audiology, Department of Clinical and Experimental Medicine, Linköping University, SE-581 82 Linköping, Sweden
e-mail: niklas.ronnberg@liu.se

Listening in noise is often perceived to be effortful. This is partly because cognitive resources are engaged in separating the target signal from background noise, leaving fewer resources for storage and processing of the content of the message in working memory. The Auditory Inference Span Test (AIST) is designed to assess listening effort by measuring the ability to maintain and process heard information. The aim of this study was to use AIST to investigate the effect of background noise types and signal-to-noise ratio (SNR) on listening effort, as a function of working memory capacity (WMC) and updating ability (UA). The AIST was administered in three types of background noise: steady-state speech-shaped noise, amplitude modulated speech-shaped noise, and unintelligible speech. Three SNRs targeting 90% speech intelligibility or better were used in each of the three noise types, giving nine different conditions. The reading span test assessed WMC, while UA was assessed with the letter memory test. Twenty young adults with normal hearing participated in the study. Results showed that AIST performance was not influenced by noise type at the same intelligibility level, but became worse with worse SNR when background noise was speech-like. Performance on AIST also decreased with increasing memory load level. Correlations between AIST performance and the cognitive measurements suggested that WMC is of more importance for listening when SNRs are worse, while UA is of more importance for listening in easier SNRs. The results indicated that in young adults with normal hearing, the effort involved in listening in noise at high intelligibility levels is independent of the noise type. However, when noise is speech-like and intelligibility decreases, listening effort increases, probably due to extra demands on cognitive resources added by the informational masking created by the speech fragments and vocal sounds in the background noise.

Keywords: speech-in-noise, cognition, working memory, updating, listening effort, cognitive spare capacity

INTRODUCTION

Speech understanding requires the interplay of top–down and bottom–up processes. Top–down processes include cognitive abilities that allow speech perception and comprehension (Davis and Johnsrude, 2007; Besser et al., 2013), while bottom–up processes include the perception of sound and the ability to hear. Hearing can be regarded as a mainly passive function that provides access to the auditory world via perception of sounds. Listening can then be viewed as a higher order function that requires intention and attention (Kiessling et al., 2003; Pichora-Fuller and Singh, 2006). Every day we hear many sounds, but we only listen to some of them. We hear the hum from the refrigerator but we may listen attentively to the news on the radio. Consequently, listening is required when heard information is to be processed for comprehension and to be remembered. However, the processes involved in listening, intention and attention, load on cognitive resources and therefore demand expenditure of effort (Kiessling et al., 2003; Pichora-Fuller and Singh, 2006).

In favorable listening conditions the speech signal is intact and understanding is implicit and automatic (Rönnberg, 2003; Rönnberg et al., 2008, 2013). However, when listening takes place in adverse conditions, a mismatch between the input from the speech signal and the phonological representations that are stored in long term memory may occur. Then explicit processing is needed for speech recognition. Thus, having a good cognitive capacity facilitates speech recognition in adverse listening conditions (Edwards, 2007; Akeroyd, 2008; Avivi-Reich et al., 2014). Adverse conditions may arise due to signal degradation caused by an unfamiliar speaker, competing background sounds, signal processing in a hearing aid, or hearing impairment (Stenfelt and Rönnberg, 2009; Mattys et al., 2012). Therefore, more cognitive resources appear to be needed when listening in noise than in quiet (Larsby et al., 2005; Pichora-Fuller and Singh, 2006; Edwards,

2007; Akeroyd, 2008; Mishra et al., 2013a; Ng et al., 2013a). Even though low levels of noise can be beneficial for speech perception of weak signals through stochastic resonance (Moss et al., 2004), for well audible and clear speech noise result in worse speech perception that load the cognitive resources. These cognitive resources may include working memory and executive functions (Rönnberg et al., 2010, 2013). Working memory is the ability to temporarily store and process information (Baddeley, 2000). During speech comprehension, executive functions are required to update working memory with new information and simultaneously remove old information (Miyake et al., 2000). It has been suggested that both working memory and updating processes are involved in disambiguating degraded speech and inferring absent information when listening takes place in adverse conditions (Rudner et al., 2011b). This may compensate for speech understanding difficulties (Rönnberg et al., 2008, 2013; Rudner et al., 2011a; Mishra et al., 2013a). However, it seems that the relation between speech perception in noise and working memory capacity (WMC) is stronger when speech is masked by a fluctuating masker compared to stationary noise (Gatehouse et al., 2003; George et al., 2007; Lunner and Sundewall-Thoren, 2007; Rudner et al., 2009, 2011a; Rönnberg et al., 2010; Koelewijn et al., 2012; Zekveld et al., 2013). An explanation for this might be that individuals with greater cognitive capacity are better able to utilize the short periods with increased signal-to-noise ratio (SNR) to infer information that is masked when the noise is louder (Duquesnoy, 1983), but they might also be better to inhibit the distracting effect of the noise.

Cognitive resources are consumed in the act of listening, which in turn leaves fewer resources to process the auditory information at a higher level (Rudner and Lunner, 2013). The residual cognitive resources after successful listening has taken place are referred to as cognitive spare capacity (Mishra et al., 2010; Rudner et al., 2011a). It has been shown that cognitive spare capacity is sensitive to processing load relating to both memory storage requirements (Mishra et al., 2013a,b) and background noise (Mishra et al., 2013a). Rönnberg et al. (2014) showed an effect of SNR with decreased memory performance in poorer SNR for individuals with normal hearing and high WMC, using the Auditory Inference Span Test (AIST). This test is designed to measure the ability to apply different levels of cognitive processing to auditory information as an objective measure of listening effort. These levels are designed to load differently on working memory and the executive function of updating. When background noise level increased the memory performance decreased, even though speech intelligibility levels were better than 90% (Rönnberg et al., 2014). This suggests that more cognitive resources were engaged in listening when background noise increased, which reduced residual resources needed to remember the auditory information. However, this was only true for individuals with greater WMC. This indicated that the test might be too difficult for individuals with less WMC, and that the extra demands the noise put on the cognitive system did not further decrease the overall low memory performance. Other studies have showed an effect of improved memory performance for hearing impaired individuals with high WMC when a noise reduction algorithm was used (Ng et al., 2013a). This suggests that background noise affects memory performance for individuals with normal hearing as well as individuals with hearing impairment, but that this effect is dependent on task difficulty as well as the individual's WMC.

Limited WMC is gradually consumed by increasing processing demands when listening takes place in adverse conditions, leaving fewer resources to process and store information (Pichora-Fuller and Singh, 2006; Schneider, 2011), or in other words, leading to less cognitive spare capacity (Rudner et al., 2011a; Rudner and Lunner, 2014). Therefore, an individual with higher WMC is likely to cope better with adverse listening conditions than an individual with lower WMC (Lunner, 2003; Larsby et al., 2005; Pichora-Fuller and Singh, 2006; Foo et al., 2007; Pichora-Fuller, 2007; Rudner et al., 2009; Schneider, 2011). When a modulated masker is used, this difference is expected to be more pronounced (Koelewijn et al., 2012; Zekveld et al., 2013). Depending on the SNR, the modulated noise can divide the speech signal into intelligible and unintelligible parts. This is because the modulated noise contains short periods where the masker has low magnitude resulting in higher SNRs, where speech recognition is better, which in turn might lead to a release from masking of the target speech (Festen and Plomp, 1990). The cognitive processes, WMC and updating ability (UA), store and update unidentified disjointed parts of the speech signal, caused by the modulated masker, in working memory until the speech information can be resolved. Consequently, an individual with greater cognitive capacity is likely to be more capable to decode speech embedded in a modulated masker and thereby better speech recognition. As processing continues, the contents of working memory are continually updated with new information and old pieces of information are discarded (Rudner et al., 2011b). Therefore, an individual with greater cognitive capacity will perform better on a task that tests storage and processing of auditory information compared to an individual with fewer cognitive resources. More specifically, in easy listening conditions with low cognitive loads, there would neither be a significant performance difference between individuals with high or low WMC, nor between individuals with high or low UA, since task demands are low. However, in adverse listening conditions or when task demands require more cognitive processes, as updating information or processing of information in working memory, individuals with higher cognitive capacity are likely to perform better. Finally, when the masker is modulated, the difference in AIST performance between individuals with high cognitive capacity and individuals with low cognitive capacity is likely to be greater than in steady-state noise (Koelewijn et al., 2012; Zekveld et al., 2013).

The aim of the present study was for the first time to test whether type of noise influences listening effort measured using the AIST (Rönnberg et al., 2011) at high speech intelligibility levels. AIST performance was expected to be best in amplitude modulated noise (AMN) compared to steady state noise (SSN) and the international speech test signal (ISTS) when intelligibility was at equal level for all noise types. We also expected AIST performance to decrease with increasing noise level, as also shown by Rönnberg et al. (2014). Furthermore, we expected that participants with better cognitive capacity, i.e., higher WMC and better UA, would show better AIST performance than participants with worse cognitive capacity, similar to Rönnberg et al. (2014). Also,

participants with high cognitive capacity were expected to perform better than participants with lower cognitive capacity on AIST tasks presented at poorer SNRs in modulated noise with high memory and processing demands.

MATERIALS AND METHODS

PARTICIPANTS

Twenty participants with normal hearing thresholds, 11 women and 9 men, with a mean age of 35 years (SD: 4.4, range 28–42) accepted to be part of the study. They were all native Swedish speakers. Baseline audiometry was done (in a sound treated room according to ISO 8253-1:2010) to verify the inclusion criteria of hearing thresholds better than or equal to 20 dB HL for the frequencies 250–4000 Hz in both ears. These frequencies were used as inclusion criteria since there is little information in the speech material used above these frequencies. Three participants did not have normal hearing for all frequencies (125–8000 Hz). One participant had a threshold of 30 dB HL at 6000 Hz at the worst ear, one participant 35 dB HL at 6000 Hz and 40 dB HL at 8000 Hz at the worse ear, and one participant 30 dB at 125 Hz at the worse ear. The participants had self-reported normal visual acuity (after correction), and no tinnitus problems. All had participated in a previous study (Rönnberg et al., 2014). The study was approved by the Regional Ethical Review Board in Linköping.

MATERIALS

The AIST test (Rönnberg et al., 2011, 2014) uses five-word matrix-type sentences in Swedish, the Hagerman sentences (Hagerman, 1982; Hagerman and Kinnefors, 1995). These sentences always have the same structure: name, verb, number, adjective, item. For example "Anna has four new gloves," see **Figure 1**. The tokens for each category are selected from a closed set of 10 items. Thus, the Hagerman sentences have low redundancy, which makes it impossible to predict any of the words from the context provided in the sentence.

Three noise types were used in the experiment. One of these was the original speech-shaped steady state noise (SSN) by Hagerman (1982) which has the same long-term average spectrum as the speech material. The second noise type (AMN) was the same as SSN but amplitude modulated with a modulation frequency of 5 Hz and a modulation depth of 20 dB. The third noise type was the ISTS (Holube et al., 2010), which consists of six voices reading a story in six different languages. These recordings are cut into 500 ms segments, which are then randomized and concatenated. This method ensures a natural speech signal that is largely non-intelligible.

The test was administered at three different SNRs targeting a speech intelligibility of above 90% but below 100%, see **Figure 2**. This ensured reasonably good speech recognition, while the noise level theoretically caused a challenging listening situation. In a previous study (Rönnberg et al., 2014), the AIST was administered in SSN at three SNRs (−2, −4, and −6 dB). These SNRs corresponded to the average speech intelligibility levels of 97, 96, and 91% in SSN. Ten participants with normal hearing, none of whom took part in the present study, were recruited to determine SNRs for the same three speech intelligibility levels: 97% (SNR1), 96% (SNR2), and 91% (SNR3) for the target sentences embedded in AMN and ISTS. Matching speech intelligibility levels between noise types enabled comparison in AIST performance between noise types, and also made for a very conservative test of differences in listening effort across noise types and SNRs. The SNRs were obtained by altering the noise level, while holding the speech level constant. The sound was presented bilaterally through headphones.

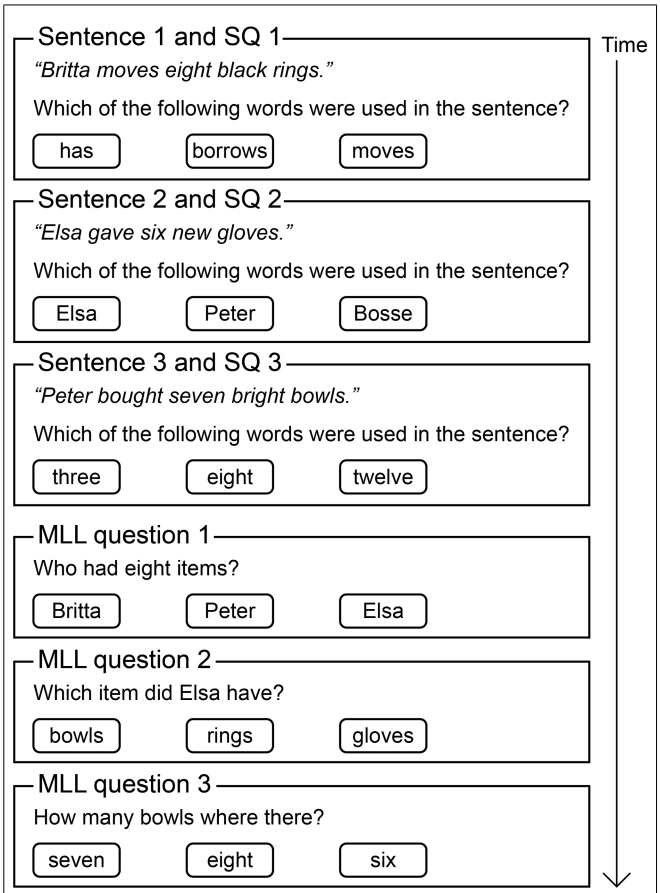


FIGURE 1 | Schematic of the Auditory Inference Span Test (AIST). A sub-list of three Hagerman sentences with SQ are shown. These are then followed by three memory load level (MLL) questions, all of these belong to the same MLL. MLL 2 questions are shown.

AUDITORY INFERENCE SPAN TEST

The AIST is a dual-task hearing-in-noise test, combining auditory and memory processing (Rönnberg et al., 2011). The participants' task is to recall and process the information from the sentences and respond in a three-alternative forced-choice procedure. In the present study, a total of nine sentences, all belonging to the same original list (Hagerman, 1982) of ten sentences, were presented consecutively in each noise type at each SNR. This was to keep speech intelligibility balanced, and to avoid duplicate answer alternatives. To verify speech recognition, one word from each sentence was probed immediately after the presentation [this will be termed sentence question (SQ)]. The accuracy and timing of the responses

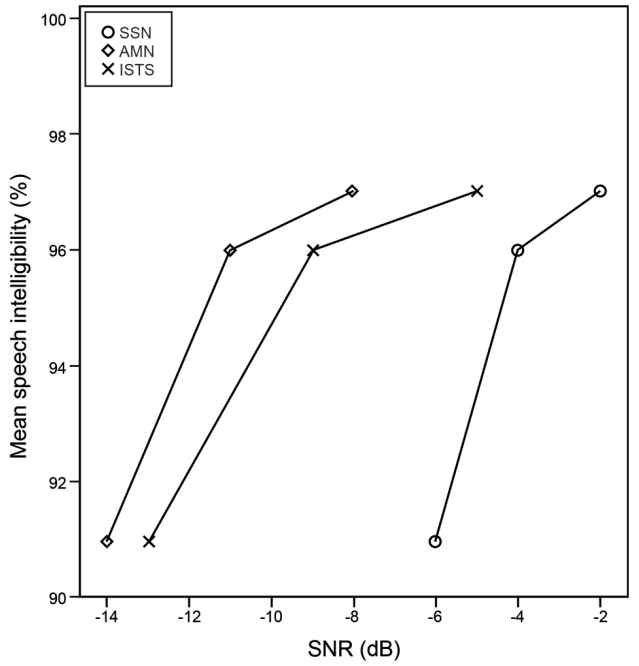


FIGURE 2 | Signal-to-noise ratio threshold for speech intelligibility levels 91, 96, and 97%.

to these questions were recorded. The AIST was administered in accordance with the standard procedure (Rönnberg et al., 2011). After each sub-list of three sentences, the participant was prompted to answer three sequentially presented multiple choice questions about the information given in the sentences, see **Figure 1**. These questions were designed to engage one of three levels of cognitive processing, called memory load levels (MLLs). Only one MLL was probed at a time, using three different questions. The multiple choice alternatives were names, numbers, or items. The order of presentation of MLLs was balanced between participants to avoid order effects. MLL 1 tapped into memory storage by asking the participant to recall which of three given words occurred in the sentences presented, e.g., "Which of the following items was used in the sentences." This type of question could be answered simply by scanning information held in working memory. MLL 2 also tapped into memory storage but also required updating, e.g., "What item did Britta have?" This type of question could be answered by scanning the sentences to find the appropriate name, updating working memory to maintain the relevant sentence and then scanning the sentence to find the relevant item. Consequently, MLL 2 made greater demands on working memory storage and updating than MLL 1. MLL 3 was the most cognitively demanding level. It required storage and updating of information in working memory, as well as processing of the information from all three sentences presented, e.g., "Which item was there most of?" This type of question could be answered by scanning the sentences for the relevant information and comparing between sentences to find the information that met the criterion. After that, memory could be updated to retain the appropriate sentence and identify the correct answer. Thus, MLL 3 made greater cognitive demands than MLL 2, specifically in terms of working memory storage, comparing characteristics and updating. Correct responses related equally often to the first, second, and third sentences and a balancing procedure ensured that this applied across conditions and participants. The AIST score was the number of questions that were correctly answered for each MLL in each noise type at each SNR.

COGNITIVE TESTS

The reading span test (RS; Rönnberg et al., 1989; Daneman and Merikle, 1996) is a well-established test of working memory (Unsworth and Engle, 2007). A short version in Swedish, with a maximum score of 28, was used in the present study (Rönnberg et al., 2014). Grammatically correct three-word sentences were presented, one word at the time, on the computer screen. Half of the sentences were reasonable and half were absurd. After each sentence, the participant was asked to judge whether it made sense or not. After each set of between 2 and 5 sentences, the participant's task was to recall in serial order either the first or the last words of each of the sentences in the set. The prompt "first" or "last" was provided only after set presentation was complete. The reading span score was the number of correctly recalled words.

The letter memory test (LM) evaluates the executive function of updating (Miyake et al., 2000). Lists of consonants were presented with capital letters one at a time on the computer screen, and the participant's task was to recall the last four letters in the correct order. The length of the lists was either 5, 7, 9, or 11 letters long, and the presentation order was randomized. Thus, list length could not be accurately predicted. The letter memory score was the number of the four target letters that were correctly recalled in serial order for each list.

SET UP AND TEST PROCEDURE

The AIST experiment was administered with an application developed in Matlab (R2013a; Rönnberg et al., 2014). Visual stimuli were presented on a 14″ computer screen, and auditory stimuli via an M-Audio FireWire 410 audio interface through a pair of Sennheiser HDA 200 headphones with the speech level calibrated to an output level of 60 dB SPL. The testing took place in a single session in a quiet room. Even if the room was not sound attenuated, the test environment was deemed quiet enough not to affect the tests conducted. Before the test started, the participants read written instructions as a complement to instructions given orally by the test supervisor. The total testing time was at most 30 min.

STATISTICAL ANALYSES

The data collected in this study were analyzed together with AIST performance in SSN as well as cognitive measurements of the participants collected in a previous study (Rönnberg et al., 2014). Repeated measures analyses of variance were performed on accuracy scores generated by the AIST. Bonferroni adjustment for multiple comparisons was applied as appropriate. To determine effects of other measurements on AIST performance, Pearson's correlation analyses were used. These analyses started with total AIST score (pooled over noise type, SNR, and MLL), then AIST performance in each noise type

(pooled over SNR and MLL), AIST performance in each SNR (pooled over noise type and MLL), and AIST performance in each MLL (pooled over noise type and SNR), and then AIST performance in each SNR in each noise type (pooled over MLL). All statistic calculations were performed using IBM SPSS Statistics 22.

RESULTS

COGNITIVE TESTS

Mean performance on the RS was 16.2 (SD = 3.7, max = 28), and mean performance on the LM was 36 (SD = 5.2, max = 48), see **Table 1**. There was no statistically significant correlation between RS and LM scores ($r = 0.25$, $p = 0.29$).

SPEECH INTELLIGIBILITY

Speech intelligibility data collected in the previous study (Rönnberg et al., 2014) were reanalyzed in the current study. A repeated measures ANOVA with one within group variable, SNR (SNR1, SNR2, SNR3) showed a significant effect of SNR [$F(2,38) = 27.5$, $p < 0.001$, $\eta_p^2 = 0.59$]. *Post hoc* test showed a significant decrease in speech intelligibility levels between SNR1 and SNR2 ($p = 0.035$), between SNR1 and SNR3 ($p < 0.001$), as well as between SNR2 and SNR3 ($p < 0.001$). Speech intelligibility data was not collected in this study and thus speech intelligibility levels for AMN as well as for ISTS are based on the equalization data obtained from 10 subjects prior to the current study.

AUDITORY INFERENCE SPAN TEST

The mean AIST performance in SSN was 16.4 (SD = 4.9) when performance was pooled over SNRs and MLLs (max = 27). In AMN the mean AIST performance was 18.1 (SD = 5.1), and in ISTS the mean AIST performance was 16.5 (SD = 4.5; see **Tables 1** and **2**; **Figure 3A**). The mean AIST performance in SNR1 was 17.6 (SD = 4.2), for SNR2 it was 17.2 (SD = 4.4), and for SNR3 it was 16.2 (SD = 4.4), when performance was pooled over noise types and MLLs (max = 27). The mean AIST performance was 21.5 (SD = 3.0) for MLL 1, 15.2 (SD = 5.8) for MLL 2, and 14.2 (SD = 5.0) for MLL 3, when performance was pooled over noise types and SNRs (see **Table 1**; **Figure 3B**).

A repeated measures ANOVA with three within group variables, noise type (SSN, AMN, ISTS), SNR (SNR1, SNR2, SNR3), and MLL (MLL 1, MLL 2, MLL 3), revealed no significant effect of noise type, a tendency to significant effect of SNR [$F(2,38) = 2.91$, $p = 0.067$, $\eta_p^2 = 0.13$], and a significant effect of MLL [$F(2,38) = 29.98$, $p < 0.001$, $\eta_p^2 = 0.61$]. *Post hoc* tests showed a significant decrease in performance between MLL 1 and MLL 2 and between MLL 1 and MLL 3 ($p < 0.001$), but there was no significant difference between MLL 2 and MLL 3 (see **Table 1**; **Figure 3B**). A significant two-way interaction between noise type and SNR was found [$F(4,76) = 2.64$, $p = 0.040$, $\eta_p^2 = 0.12$; see **Tables 1** and **2**; **Figure 3C**]. Analyses of simple main effects revealed no differences in AIST performance between SNRs for SSN or for AMN, but for ISTS [$F(2,38) = 10.01$, $p < 0.001$, $\eta_p^2 = 0.35$]. *Post hoc* tests showed a significant decrease in memory performance on AIST between SNR1 and SNR2 ($p = 0.026$) as well as between SNR1 and SNR3 ($p = 0.002$), but not between SNR2 and SNR3. There were no other significant interactions.

Table 1 | Mean scores and SDs in parenthesis, for the cognitive tests and factorwise Auditory Inference Span Test (AIST) performance.

Cognitive tests			
Reading span score	16.2 (3.7), range 11–23, max 28		
Letter memory score	36.5 (5.2), range 23–46, max 48		
AIST performance			
Noise type (max = 27)	**Mean**	**95% Confidence Interval**	
		Lower Bound	**Upper Bound**
SSN	16.45 (4.9)	14.13	18.76
AMN	18.15 (5.1)	15.77	20.53
ISTS	16.50 (4.5)	14.41	18.59
SNR (max = 27)	**Mean**	**95% Confidence Interval**	
		Lower Bound	**Upper Bound**
SNR1	17.65 (4.2)	15.66	§19.62
SNR2	17.25 (4.4)	15.18	19.32
SNR3	16.20 (4.2)	14.16	18.24
Noise type and SNR (max = 9)	**Mean**	**95% Confidence Interval**	
		Lower Bound	**Upper Bound**
SSN SNR1	5.40 (2.0)	4.46	6.34
SNR2	5.75 (2.1)	4.74	6.76
SNR3	5.30 (1.9)	4.39	6.21
AMN SNR1	5.90 (1.8)	5.07	6.73
SNR2	6.20 (1.9)	5.32	7.08
SNR3	6.05 (2.1)	5.08	7.02
ISTS SNR1	6.35 (1.7)	5.54	7.16
SNR2	5.30 (1.8)	4.45	6.15
SNR3	4.85 (1.7)	4.07	5.63
MLL (max = 27)	**Mean**	**95% Confidence Interval**	
		Lower Bound	**Upper Bound**
MLL 1	21.50 (3.0)	20.08	22.92
MLL 2	15.20 (5.8)	12.47	17.93
MLL 3	14.25 (5.0)	11.89	16.60

AIST performance and reading span score

A significant positive correlation was found between total AIST performance and reading span score ($r = 0.712$, $p < 0.001$), showing that a higher reading span score was associated with better

Table 2 | Mean AIST performance for each SNR in each noise type pooled over MLLs.

	SSN	AMN	ISTS	Mean
SNR1	5.40	5.90	6.35	5.88
SNR2	5.75	6.20	5.30	5.75
SNR3	5.30	6.05	4.85	5.40
Mean	5.48	6.05	5.50	

general AIST performance (see **Table 3**). As shown in **Table 3**, reading span score also correlated positively with AIST performance in all three noise types, in all three SNRs, as well as with all three MLLs. More specifically in SSN, reading span score correlated with AIST performance in SNR1. In the modulated noise types (AMN and ISTS), reading span score correlated with AIST performance in SNR2 as well as SNR3.

AIST performance and letter memory score

Letter memory score did not significantly correlate with total AIST performance (see **Table 3**). The only significant correlation between Letter memory score and AIST performance was found between Letter memory score and AIST performance in SNR1 ($r = 0.495$, $p < 0.05$). As shown in **Table 3**, Letter memory score correlated with AIST performance in SNR1 for the modulated noise types (AMN and ISTS).

Sentence questions

When SQ performance was pooled over SNRs the mean score was 26.8 (SD = 0.4) in SSN, in AMN the mean score was 26.8 (SD = 0.5), and in ISTS it was 25.7 (SD = 1.4), maximum score was 27, see **Table 4** and **Figure 4A**. A repeated measures ANOVA with two within group variables, noise type (SSN, AMN, ISTS) and SNR (SNR1, SNR2, SNR3), showed a significant effect of noise type [$F(2,38) = 12.79$, $p < 0.001$, $\eta_p^2 = 0.40$], but there was only a tendency toward significant effect of SNR [$F(2,38) = 2.59$, $p = 0.088$, $\eta_p^2 = 0.12$]. *Post hoc* tests revealed a significantly better SQ performance in SSN than in ISTS ($p = 0.006$), as well as in AMN compared to in ISTS ($p = 0.004$), but there was no significant difference in SQ performance between SSN and AMN. A significant two-way interaction between noise type and SNR was found [$F(4,76) = 2.96$, $p = 0.025$, $\eta_p^2 = 0.14$]. Analyses of simple main effects revealed significant differences in SQ performance between SNRs for ISTS [$F(2,38) = 3.35$, $p = 0.046$, $\eta_p^2 = 0.15$], but only a tendency toward significant effect for SSN [$F(2,38) = 2.84$, $p = 0.071$, $\eta_p^2 = 0.13$] and no effect for AMN. *Post hoc* tests showed a significant decrease in SQ performance in ISTS between SNR1 and SNR3 ($p = 0.047$), as well as a tendency toward significant difference between SNR1 and SNR2 ($p = 0.074$), but there was no significant difference between SNR2 and SNR3. Performance on SQs did not significantly correlate with WMC or with UA.

When response times, see **Table 4** and **Figure 4B**, was assessed in a repeated measures ANOVA with two within group variables, noise type (SSN, AMN, ISTS), SNR (SNR1, SNR2, SNR3), a significant effect of noise type [$F(2,38) = 5.48$, $p = 0.008$, $\eta_p^2 = 0.23$] was revealed as well as a significant effect of SNR [$F(2,38) = 5.94$, $p = 0.006$, $\eta_p^2 = 0.24$]. *Post hoc* tests showed a significant increase in response time between SSN and ISTS ($p = 0.045$), but there were no significant differences between SSN and AMN, or between AMN and ISTS. *Post hoc* tests also showed a significant increase in response time between SNR1 and SNR3 ($p = 0.010$), but there were no significant differences between SNR1 and SNR2, or between

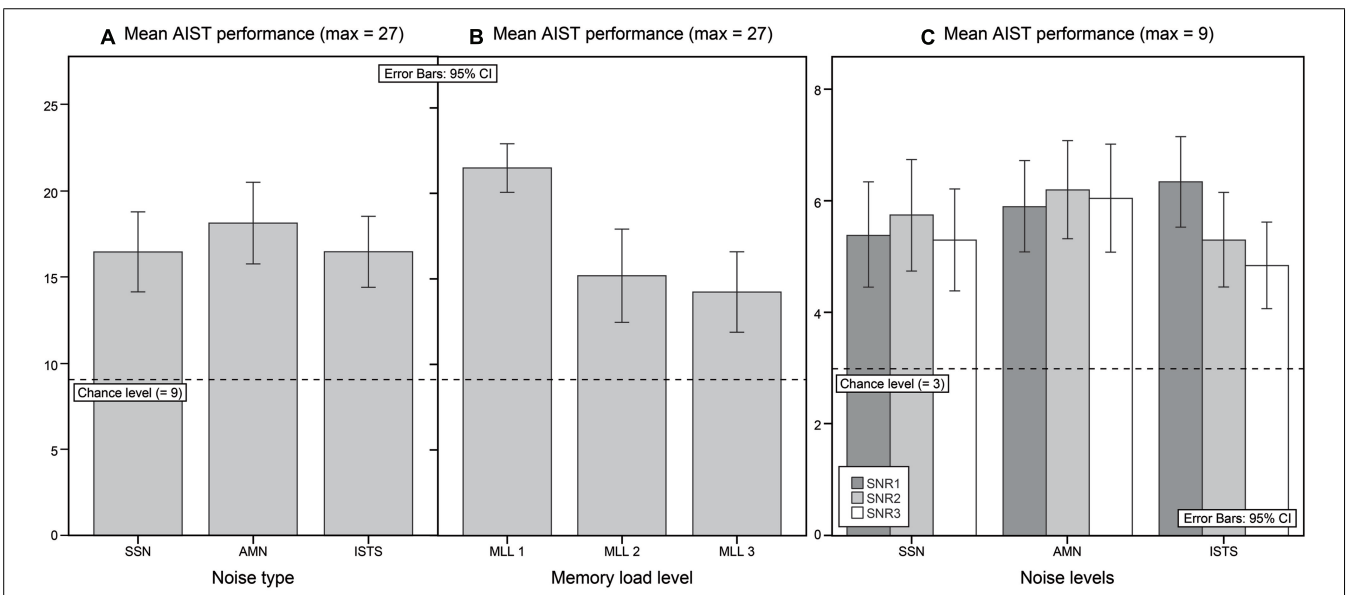


FIGURE 3 | (A) Mean AIST performance in each noise type pooled over SNRs and MLLs. The maximum score was 27. Chance level was at 9. **(B)** Mean AIST performance for each MLL pooled over noise types and SNRs. The maximum score was 27. Chance level was at 9. **(C)** Mean AIST performance in each noise type and in each SNR pooled over MLLs. The maximum score was 9. Chance level was at 3.

Table 3 | The table shows correlations between total and factorwise AIST performance and cognitive measurements (WMC and UA).

Measure		WMC	UA
AIST	Total AIST	0.712**	0.319
	Total SSN	0.460*	0.199
	Total AMN	0.616**	0.210
	Total ISTS	0.623**	0.391
	Total SNR1	0.603**	0.495*
	Total SNR2	0.569**	0.237
	Total SNR3	0.715**	0.149
	Total MLL1	0.495*	0.186
	Total MLL2	0.638**	0.374
	Total MLL3	0.656**	0.214
	SSN SNR1	0.637**	0.185
	SSN SNR2	0.108	0.200
	SSN SNR3	0.391	0.093
	AMN SNR1	0.389	0.477*
	AMN SNR2	0.636**	0.118
	AMN SNR3	0.605**	0.000
	ISTS SNR1	0.340	0.512*
	ISTS SNR2	0.602**	0.218
	ISTS SNR3	0.665**	0.283

*$p < 0.05$, **$p < 0.01$.

SNR2 and SNR3. Response time on SQs correlated positively with WMC ($r = 0.683$, $p = 0.001$) indicating that having a greater WMC yielded in a longer response time. There was no correlation found between UA and response time on SQs.

Table 4 | The table shows mean values, with standard deviations in parenthesis, for performance in each noise type in each SNR on SQ (max = 9). As well as, mean response time for each noise type in each SNR in seconds.

AIST sentence questions

Performance (max = 9)		Mean (SD)	Minimum	Maximum
SSN	SNR1	8.95 (0.2)	8.00	9.00
	SNR2	9.00 (0.0)	9.00	9.00
	SNR3	8.80 (0.4)	8.00	9.00
AMN	SNR1	8.90 (0.4)	7.00	9.00
	SNR2	9.00 (0.0)	9.00	9.00
	SNR3	8.90 (0.3)	8.00	9.00
ISTS	SNR1	8.85 (0.4)	8.00	9.00
	SNR2	8.40 (0.7)	7.00	9.00
	SNR3	8.40 (0.9)	6.00	9.00

Response time (seconds)		Mean	95% Confidence Interval	
			Lower Bound	Upper Bound
SSN	SNR1	2.1 (0.16)	1.8	2.4
	SNR2	2.2 (0.16)	1.9	2.6
	SNR3	2.4 (0.16)	2.1	2.8
AMN	SNR1	2.4 (0.21)	2.1	2.8
	SNR2	2.7 (0.23)	2.2	3.2
	SNR3	2.8 (0.23)	2.2	3.2
ISTS	SNR1	2.5 (0.16)	2.2	2.8
	SNR2	3.0 (0.34)	2.2	3.7
	SNR3	3.0 (0.30)	2.4	3.6

DISCUSSION

SPEECH INTELLIGIBILITY

Speech intelligibility levels in SSN in the present study were identified in a larger study cohort (Rönnberg et al., 2014). The speech intelligibility levels in AMN and ISTS were matched to the speech intelligibility levels in SSN prior to the study to provide equal intelligibility levels between noise types. Even though performance on SQ is not a measure of speech intelligibility, it is nevertheless an indication of how well the participant has heard the sentence. The accuracy on SQs supported the estimated speech intelligibility levels used.

AUDITORY INFERENCE SPAN TEST

Noise types

It was hypothesized that the average AIST performance would differ between noise types, even though mean speech intelligibility levels were held constant. The poorest AIST performance was expected to be found in SSN, while the best AIST performance was expected to be found in AMN. However, contrary to expectations there were no statistical significant differences in memory performance between the noise types (see **Figure 3C**). Mishra et al. (2013a) showed an increased cognitive spare capacity, as measured by improved memory performance, in ISTS compared to SSN, using lists of numbers between 13 and 99 as targets. This was not the case in the present study. The reason for this might be that the vocal sounds and speech fragments add an additional informational masking interfering more with the speech information in the sentences compared to the numbers used by Mishra et al. (2013a). This in turn would add more demands on the cognitive system leading to less cognitive spare capacity. The AMN contains short periods with less noise which might make it possible to achieve the same speech intelligibility level as for SSN but with less cognitive demands (Duquesnoy, 1983), but there was no statistical significant improved memory performance in AMN compared to SSN or ISTS (see **Figure 3C**). This suggests that for young adults with normal hearing, in SNRs targeting 90% speech intelligibility or better, the type of noise is not of importance for memory performance of the information in the sentences.

Signal-to-noise ratio

Speech intelligibility levels were matched between all noise types at SNR1, as well as at SNR2 and at SNR3 (see **Figure 2**). Therefore, the amount of amplitude change of the noise between SNR1 and

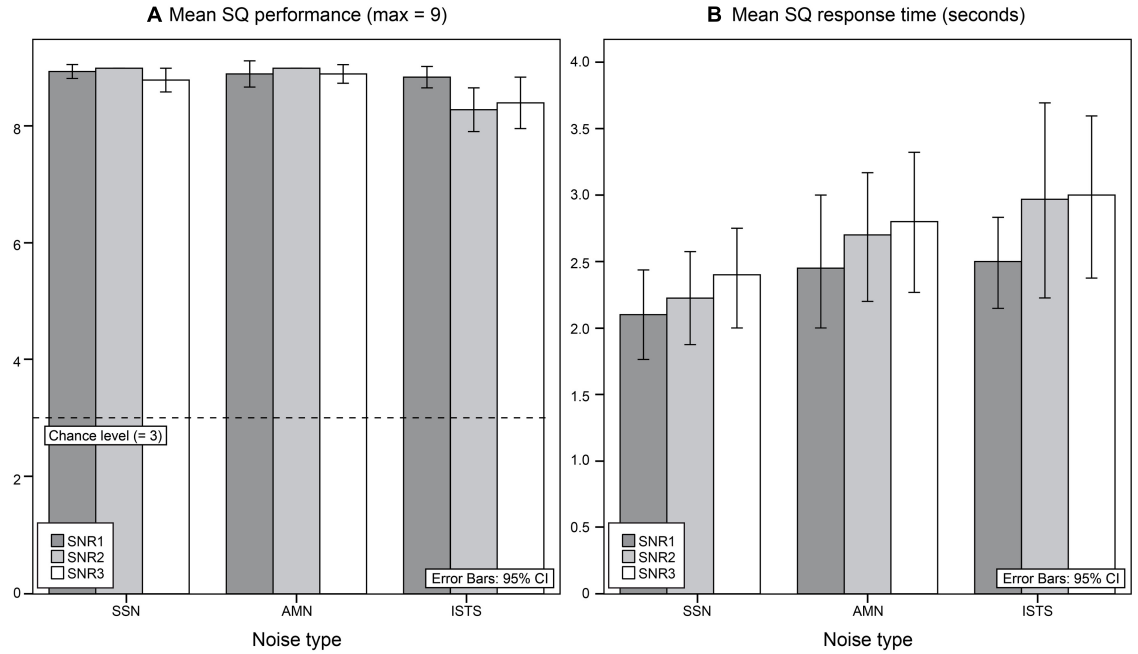


FIGURE 4 | (A) Mean sentence question (SQ) performance for SNR type in each noise type. The maximum score was 9. Chance level was at 3. **(B)** Mean response time (in seconds) for SQ questions for each SNR in each noise type.

SNR2, as well as between SNR2 and SNR3, differed between noise types, i.e., SNR1 was different in different noise types but corresponded to the same speech intelligibility level (see **Figure 2**). Access to the information in the sentences is essential for accurate AIST performance. Since all SNRs gave a mean speech intelligibility level of 90% or better, access to the information was not appreciably limited at any of the SNRs (see **Figure 2**).

Based on the previous study (Rönnberg et al., 2014), we hypothesized that a decreased SNR would force an increase in cognitive processing of auditory information, leading to less cognitive spare capacity resulting in reduced AIST performance. The tendency toward a statistically significant effect of SNR on AIST performance (see **Tables 1** and **2**; **Figure 3C**) suggested that the cognitive spare capacity, as measured by memory performance on AIST, was reduced by increasing noise level. Similar results have also been found in other studies (Mishra et al., 2013a; Ng et al., 2013a,b; Rönnberg et al., 2014). However, in the present study, increasing noise level only reduced AIST performance when ISTS was used as background noise. This suggests that increasing background noise at the high intelligibility levels used in the present study only influences listening effort when noise is speech-like (see **Figure 3C**).

When listening in AMN, young adults with normal hearing are likely to be able to utilize the short periods with increased SNR to infer information that is masked when the noise level is louder (Duquesnoy, 1983) which would give rise to release from masking (Festen and Plomp, 1990). As a result, the decrease in SNR for AMN might not be particularly more demanding when listening in SNRs targeting 90% speech intelligibility or better. Nevertheless, for ISTS, the noise level seemed to have an impact on the cognitive processes involved leading to less cognitive spare capacity and decreased memory performance on AIST (see **Tables 1** and **2**; **Figure 3C**). Even if the ISTS is largely non-intelligible (Holube et al., 2010), the voices and speech fragments in ISTS may promote informational masking (Francart et al., 2011) which would add to the cognitive load since ISTS will interfere with the Hagerman sentences at different linguistic levels (Tun et al., 2002; Brouwer et al., 2012). Consequently, since ISTS adds more cognitive load, AIST performance in ISTS is more sensitive to decreased SNR than in the other noise types. As a result, the decrease in AIST performance with worse SNR in ISTS cannot be explained by reduced intelligibility alone since SNR did not significantly affect AIST performance in SSN or in AMN.

Interestingly, the correlations with WMC, i.e., reading span score, indicated that WMC had an impact on performance in AIST when presentation took place in SSN with SNR1, but not with the other SNRs (see **Table 3**). A reason for this might be that SSN masks the signal at worse SNRs, and when the signal becomes inaudible, a greater WMC does not improve speech intelligibility. On the other hand, when SNR is better and the signal is only partly masked by the SSN, a greater WMC might facilitate speech intelligibility by storing partly heard sounds of the speech signal until these can be disambiguated. The relation between speech recognition in noise and WMC is more evident in modulated noise where individuals with high WMC have better speech recognition in noise performance compared to individuals with less WMC (Gatehouse et al., 2003; George et al., 2007; Zekveld et al., 2013), which might also explain the relation between WMC and AIST performance in SSN. For the modulated noise, WMC was of importance for memory

performance when the SNR was more demanding (see **Table 3**). This suggests that when listening takes place in more troublesome listening conditions, such as increased SNR and modulated noise, WMC is more occupied with listening, and individuals with higher cognitive capacity are likely to have more cognitive spare capacity after listening and consequently perform better on the memory task than individuals with less cognitive capacity. Consequently, individuals with greater cognitive capacity will probably experience less listening effort than individuals with less cognitive capacity. On the other hand, when listening takes place in modulated noise in SNR1, the listening condition might be described as fairly simple which explains why, the extra WMC capacity did not add an additional advantage.

Another way to explain the correlations between AIST performance and WMC is based on attention. One may expect that a person with a higher WMC is better to filter out the desired signal (speech) and suppress the unwanted signal (noise) than a person with worse WMC. There are indications of such mechanisms in the literature. In an auditory brainstem response measurement it was found that the neural amplitude increased when focusing on the signal and decreased when adding a cognitive load (distractor; Sorqvist et al., 2012). This modulation of the neural response was correlated with the persons WMC. Other studies have indicated that attention and WMC correlates with spatial speech recognition performance in adults (Neher et al., 2011) and that attention supports language processing in children (Astheimer et al., 2014). However, there are other studies that have found correlation between WMC and speech perception that is unrelated to attention skills (Tamati et al., 2013). The current study did not measure attention *per se*, but it is very plausible that a better WMC facilitated auditory attentional filtering of the sentence and thereby improved both speech recognition and ability to store the information crucial for AIST performance.

Updating ability, i.e., Letter memory score, did not correlate with total AIST performance (see **Table 3**). However, having a greater UA improved AIST performance in SNR1, more specifically for SNR1 in the modulated noise types (AMN and ISTS) but not in SSN. This is consistent with the previous study where no interactions were found between AIST performance and SNRs when UA was used as a between-group variable and SSN was used as masker (Rönnberg et al., 2014). In the modulated noise types, at the best SNR, listening is fairly undemanding why having a higher UA facilitates performance on AIST. However, when the SNR gets worse, there was no effect of UA on AIST performance. Nevertheless, there was an effect of WMC on AIST performance in worse SNRs suggesting that in more troublesome listening conditions WMC is of more importance for listening than UA. WMC improves memory performance in SSN in the easiest SNR, but UA does not improve memory performance. However, in modulated noise, WMC facilitates memory performance in the worst SNR, while UA improves memory performance in the best SNR.

Memory load level

Auditory Inference Span Test accuracy was, as expected, a function of MLL (see **Table 1**; **Figure 3B**), where performance decreased with increasing level of memory load (Mishra et al., 2013a,b; Rönnberg et al., 2014). As in the previous study (Rönnberg et al., 2014), there were no significant difference in performance on MLL2 and MLL3. Even though performance at MLL2 and MLL3 is low, performance on both MLLs are clearly above chance level. The results suggested that regardless of MLL, WMC improves memory performance on AIST. A similar effect was found in a previous study (Rönnberg et al., 2014). Also, in the previous study (Rönnberg et al., 2014) an interaction between MLL and UA showed a benefit of high UA on questions demanding more updating of information, i.e., MLL 2. This relation was not found to be significant in the present study (see **Table 3**).

Response time

Response times on MLL questions were registered in the AIST process. These response times on MLL questions were not included in the analyses. The reason for this was that the measure of response time started when the question was presented on the computer screen and continued until an answer had been given, and the test had continued to the next question. Consequently, the time it took to read and comprehend the question was part of the measured response time. However, there is a difference in the complexity of the questions, why differences in response time might be due to differences in the amount of time it took to read and comprehend the question. Nevertheless, response times on MLL questions might be analyzed when pooled over the three MLLs. It was expected that response times then would be dependent on SNRs and noise types. However, no statistically significant effect of SNR or of noise type was not found. Pooled response times on MLL questions did not change with listening conditions. Consequently, response time on AIST was not deemed to be a useful measure.

SENTENCE QUESTIONS

Performance on SQs decreased in ISTS compared to SSN and AMN, and there was an effect of SNR in ISTS but not in SSN or AMN, see **Figure 4A**. Since SQ might be considered a measure of speech recognition in the sense that the question probes that the sentence was heard, even if the three-choice procedure facilitates performance by giving possible answer alternatives as well as having a chance level of 33%, the results suggested that the general speech intelligibility levels were at the expected levels above 91% (Rönnberg et al., 2014). However, the effect of SNR only found in ISTS might suggest that speech intelligibility levels were not perfectly matched between noise types. Nevertheless, the results might also imply that speech-shaped noise in these rather favorable SNRs did not load the cognitive system to such a degree as the vocal sounds and speech fragments in ISTS did, and consequently there was no effect of SNRs for SSN and AMN on SQ accuracy. Even if ISTS is largely non-intelligible (Holube et al., 2010), it may cause additional informational masking (Francart et al., 2011) and consequently add to the cognitive load since the masker interferes with the speech material at different linguistic levels (Tun et al., 2002; Brouwer et al., 2012).

The analyses of SQ response times were based on response times correct answers as well as for incorrect answers, as there was no statistically significant difference in response time between correct and incorrect answers. Response time on SQs was an effect of noise type, with longer response times in ISTS compared to

SSN and AMN. There was also an effect of SNR with increasing response times in SNR3 compared to SNR1, see **Figure 4B**. The results suggest that more processing was needed in the more problematic listening conditions (in ISTS compared to SSN, and in SNR3 compared to SNR1) and that this processing takes longer, with longer response times as a result. It seems likely to assume that the longer response time is a measure of listening effort. SQ response time correlated with WMC and not with UA. Contrary to expectations that having a greater WMC would imply faster access time to information stored in working memory and a shorter time to retrieve the position of the correct answer alternative, instead the results showed that greater WMC rather meant longer response times. The results suggested that individuals with greater WMC spent more time reading the answer alternatives and pondering the answer; however, they did not gain from this extra time spent when considering accuracy on SQ questions. Also, having a higher WMC implies having more information held in working memory, resulting in more information to scan which would require a longer time to find the matching answer.

THE COGNITIVE MEASUREMENTS

Both the RS and the LM are delivered in visual modality, unlike the AIST which is delivered in auditory modality with visually presented multiple choice responses. This is a strength of the study, since the measurements of WMC and of UA are independent on the individual's hearing status. Furthermore, the AIST is intended to be used in the hearing aid fitting process to assess listening effort, then it is of even greater importance that the measurement of the individual's cognitive capacity is unaffected by the hearing status.

CLINICAL IMPLICATION

Performance on AIST can be expected to be lower for individuals with hearing impairment than for individuals with normal hearing. A hearing impairment decreases the signal fidelity (Plomp, 1978; Pichora-Fuller and Singh, 2006), which in turn increases the cognitive involvement in listening and consequently leaves less cognitive capacity for memory storage (Rudner et al., 2011b; Picou et al., 2013) which would be measurable with the AIST. It is well established that successful hearing aid fitting needs to take individual differences in cognitive capacity into account (Lunner et al., 2009). Hitherto, cognitive measures such as reading span have been used to demonstrate associations with ability to repeat and recall speech. The advantage of a test such as AIST is that it has the potential to measure the listening effort expended by the individual under different sets of listening conditions in which noise types, SNR and potentially hearing aid settings can be manipulated. This will allow better hearing aid fitting in the future and provides an important tool for the development of better hearing aids.

CONCLUSION

The results suggest that for young adults with normal hearing the cognitive spare capacity is reduced when background noise consists of voices and the SNR decreases. However, when speech intelligibility levels are kept constant, different masker types do not have different effects on cognitive spare capacity, at least not for intelligibility levels above 90%.

ACKNOWLEDGMENT

This work was supported by the Oticon Foundation.

REFERENCES

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

Astheimer, L., Janus, M., Moreno, S., and Bialystok, E. (2014). Electrophysiological measures of attention during speech perception predict metalinguistic skills in children. *Dev. Cogn. Neurosci.* 7, 1–12. doi: 10.1016/j.dcn.2013.10.005

Avivi-Reich, M., Daneman, M., and Schneider, B. A. (2014). How age and linguistic competence alter the interplay of perceptual and cognitive factors when listening to conversations in a noisy environment. *Front. Syst. Neurosci.* 8:21. doi: 10.3389/fnsys.2014.00021

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise–a review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459

Brouwer, S., Van Engen, K. J., Calandruccio, L., and Bradlow, A. R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: language familiarity and semantic content. *J. Acoust. Soc. Am.* 131, 1449–1464. doi: 10.1121/1.3675943

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546

Davis, M. H., and Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hear. Res.* 229, 132–147. doi: 10.1016/j.heares.2007.01.014

Duquesnoy, A. J. (1983). Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *J. Acoust. Soc. Am.* 74, 739–743. doi: 10.1121/1.389859

Edwards, B. (2007). The future of hearing aid technology. *Trends Amplif.* 11, 31–45. doi: 10.1177/1084713806298004

Festen, J. M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 74, 739–743. doi: 10.11121/1.400247

Foo, C., Rudner, M., Rönnberg, J., and Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requeres explicit cognitive storage and processing capacity. *J. Am. Acad. Audiol.* 18, 618–631. doi: 10.3766/jaaa.18.7.8

Francart, T., van Wieringen, A., and Wouters, J. (2011). Comparison of fluctuating maskers for speech recognition tests. *Int. J. Audiol.* 50, 2–13. doi: 10.3109/14992027.2010.505582

Gatehouse, S., Naylor, G., and Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Int. J. Audiol.* 42, 77–85.

George, E. L. J., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., and Houtgast, T. (2007). Auditory and nonauditory factors affecting speech reception in noise by older listeners. *J. Am. Acad. Audiol.* 121, 2362–2375. doi: 10.1121/1.2642072

Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scand. J. Audiol.* 11, 79–87. doi: 10.3109/01050398209076203

Hagerman, B., and Kinnefors, C. (1995). Efficient Adaptive Methods for Measuring Speech Reception Threshold in Quiet and in Noise. *Scand. J. Audiol.* 24, 71–77. doi: 10.3109/01050399509042213

Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *Int. J. Audiol.* 49, 891–903. doi: 10.3109/14992027.2010.506889

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T. H., et al. (2003). Candidaturefor and delivery of audiological services – special needs of older people. *Int. J. Audiol.* 42, 92–101. doi: 10.3109/14992020309074650

Koelewijn, T., Zekveld, A. A., Festen, J. M., and Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a singletalker masker. *Ear Hear.* 32, 291–300. doi: 10.109/aud.0b013e3182310019

Larsby, B., Hällgren, M., Lyxell, B., and Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *Int. J. Audiol.* 44, 131–143. doi: 10.1080/14992020500057244

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42, 49–58. doi: 10.3109/14992020309074624

Lunner, T., Rudner, M., and Rönnberg, J. (2009). Cognition and hearing aids. *Scand. J. Psychol.* 50, 395–403. doi: 10.1111/j.1467-9450.2009.00742.x

Lunner, T., and Sundewall-Thoren, E. (2007). Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J. Am. Acad. Audiol.* 18, 604–617. doi: 10.1111/j.1467-9450.2009.00742.x

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006

Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., and Rudner, M. (2013a). Seeing the talker's face supports executive processing of speech in steady state noise. *Front. Syst. Neurosci.* 7:96. doi: 10.3389/fnsys.2013.00096

Mishra, S., Lunner, T., Stenfelt, S., Rönnberg, J., and Rudner, M. (2013b). Visual information can hinder working memory processing of speech. *J. Speech Lang. Hear. Res.* 56, 1120–1132. doi: 10.1044/1092-4388(2012/12-0033

Mishra, S., Rudner, M., Lunner, T., Stenfelt, S., and Rönnberg, J. (2010). "Speech understanding and cognitive spare capacity," in *Binaural Processing and Spatial Hearing*, eds J. M. Buchholz, T. Dau, J. Christensen-Dalsgaard, and T. Poulsen (Elsinore: ISSAR), 305–313.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moss, F., Ward, L. M., and Sannita, W. G. (2004). Stochastic resonance and sensory information processing: a tutorial and review of application. *Clin. Neurophysiol.* 115, 267–281. doi: 10.1016/j.clinph.2003.09.014

Neher, T., Laugesen, S., Jensen, N. S., and Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *J. Acoust. Soc. Am.* 130, 1542–1558. doi: 10.1121/1.3608122

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013a). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.* 52, 433–441.

Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013b). Improved cognitive processing of speech for hearing imparied users with noise reduction. *Int. J. Audiol.* 52, 433–441. doi: 10.3109/14992027.2013.776181

Pichora-Fuller, M. K. (2007). Audition and cognition: what audiologists need to know about listening. *Paper presented at the Adult Conference*, Stäfa.

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Picou, E. M., Ricketts, T. A., and Hornsby, B. W. (2013). How hearing aids, background noise, and visual cues influence objective listening effort. *Ear Hear.* 34, e52–e64. doi: 10.1097/AUD.0b013e31827f0431

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* 63, 533–549. doi: 10.1121/1.381753

Rönnberg, J. (2003). Cognition in the hearing imparied and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42, 68–76. doi: 10.3109/14992020309074626

Rönnberg, J., Arlinger, S., Lyxell, B., and Kinnefors, C. (1989). Visual evoked potentials: Relation to adult speechreading and cognitive function. *J. Speech Hear. Res.* 32, 725–735. doi: 10.1044/jshr.3204.725

Rönnberg, J., Lunner, T., Zekveld, A. A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S99–S105. doi: 10.1080/14992020802301167

Rönnberg, N., Rudner, M., Lunner, T., and Stenfelt, S. (2014). Assessing listening effort by measuring short-term memory storage and processing of speech in noise. *Speech Lang. Hear.* 17, 123–132. doi: 10.1179/2050572813Y.0000000033

Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12, 263–269. doi: 10.4103/1463-1741.70505

Rönnberg, N., Stenfelt, S., and Rudner, M. (2011). Testing listening effort for speech comprehension using the individuals' cognitive spare capacity. *Audiol. Res.* 1:e22. doi: 10.4081/audiores.2011.e22

Rudner, M., Foo, C., Rönnberg, J., and Lunner, T. (2009). Cognition and aided speech recognition in noise: specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scand. J. Psychol.* 50, 405–418. doi: 10.1111/j.1467-9450.2009.00745.x

Rudner, M., and Lunner, T. (2013). Cognitive spare capacity as a window on hearing aid benefit. *Semin. Hear.* 34, 297–306. doi: 10.1055/s-0033-1356642

Rudner, M., and Lunner, T. (2014). Cognitive spare capacity and speech communication: a narrative overview. *Biomed Res. Int.* 2014:869726. doi: 10.1155/2014/869726

Rudner, M., Ng, E. H. N., Rönnberg, N., Mishra, S., Rönnberg, J., Lunner, T., et al. (2011a). Cognitive spare capacity as a measure of listening effort. *J. Hear. Sci.* 11, 47–49. doi: 10.1155/2014/869726

Rudner, M., Rönnberg, J., and Lunner, T. (2011b). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4

Schneider, B. A. (2011). How age affects auditory-cognitive interactions in speech comprehension. *Audiol. Res.* 1:e10. doi: 10.4081/audiores.2011.e10

Sorqvist, P., Stenfelt, S., and Ronnberg, J. (2012). Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: toward a unified view of attention. *J. Cogn. Neurosci.* 24, 2147–2154. doi: 10.1162/jocn_a_00275

Stenfelt, S., and Rönnberg, J. (2009). The signal-cognition interface: interactions between degraded auditory signals and cognitive processes. *Scand. J. Psychol.* 50, 385–393. doi: 10.1111/j.1467-9450.2009.00748.x

Tamati, T. N., Gilbert, J. L., and Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: a first report. *J. Am. Acad. Audiol.* 24, 616–634. doi: 10.3766/jaaa.24.7.10

Tun, P. A., O'Kane, G., and Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychol. Aging* 17, 453–467. doi: 10.1037//0882-7974.17.3.453

Unsworth, N., and Engle, R. W. (2007). On the division of short-term and working memory: an examination of simple and complex span and their relation to higher order abilities. *Psychol. Bull.* 133, 1038–1066. doi: 10.1037/0033-2909.133.6.1038

Zekveld, A. A., Rudner, M., Johnsrude, I. S., and Rönnberg, J. (2013). The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *J. Acoust. Soc. Am.* 134, 2225–2234. doi: 10.1121/1.4817926

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 July 2014; accepted: 03 December 2014; published online: 22 December 2014.

*Citation: Rönnberg N, Rudner M, Lunner T and Stenfelt S (2014) Memory performance on the Auditory Inference Span Test is independent of background noise type for young adults with normal hearing at high speech intelligibility. Front. Psychol. **5**:1490. doi: 10.3389/fpsyg.2014.01490*

This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology.

Copyright © 2014 Rönnberg, Rudner, Lunner and Stenfelt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 20 April 2015
doi: 10.3389/fnins.2015.00124

Costs of switching auditory spatial attention in following conversational turn-taking

Gaven Lin and Simon Carlile*

Auditory Neuroscience Laboratory, Department of Physiology, School of Medical Sciences, University of Sydney, Sydney, NSW, Australia

Following a multi-talker conversation relies on the ability to rapidly and efficiently shift the focus of spatial attention from one talker to another. The current study investigated the listening costs associated with shifts in spatial attention during conversational turn-taking in 16 normally-hearing listeners using a novel sentence recall task. Three pairs of syntactically fixed but semantically unpredictable matrix sentences, recorded from a single male talker, were presented concurrently through an array of three loudspeakers (directly ahead and $+/-30^{\circ}$ azimuth). Subjects attended to one spatial location, cued by a tone, and followed the target conversation from one sentence to the next using the call-sign at the beginning of each sentence. Subjects were required to report the last three words of each sentence (speech recall task) or answer multiple choice questions related to the target material (speech comprehension task). The reading span test, attention network test, and trail making test were also administered to assess working memory, attentional control, and executive function. There was a $10.7 \pm 1.3\%$ decrease in word recall, a pronounced primacy effect, and a rise in masker confusion errors and word omissions when the target switched location between sentences. Switching costs were independent of the location, direction, and angular size of the spatial shift but did appear to be load dependent and only significant for complex questions requiring multiple cognitive operations. Reading span scores were positively correlated with total words recalled, and negatively correlated with switching costs and word omissions. Task switching speed (Trail-B time) was also significantly correlated with recall accuracy. Overall, this study highlights (i) the listening costs associated with shifts in spatial attention and (ii) the important role of working memory in maintaining goal relevant information and extracting meaning from dynamic multi-talker conversations.

Keywords: spatial attention, speech, cocktail party, switch costs, working memory, cognitive load

OPEN ACCESS

Edited by:
Mary Rudner, Linköping University, Sweden

Reviewed by:
Rachel Jane Ellis, Linköping University, Sweden
Thomas Koelewijn, VU University Medical Center, Netherlands

****Correspondence:***
Gaven Lin, Auditory Neuroscience Laboratory, Department of Physiology, University of Sydney, Anderson Stuart Building, Sydney, NSW 2006, Australia
gavenlin@gmail.com

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience

Received: *31 January 2015*
Accepted: *26 March 2015*
Published: *20 April 2015*

Citation:
Lin G and Carlile S (2015) Costs of switching auditory spatial attention in following conversational turn-taking. Front. Neurosci. 9:124. doi: 10.3389/fnins.2015.00124

Introduction

In a cocktail party environment, listeners are faced with the challenging task of separating multiple simultaneous talkers overlapping in time, frequency, and space. The auditory system is able to parse this complex mixture into meaningful perceptual objects (Griffiths and Warren, 2004) using perceived differences in spatial location (e.g., Freyman et al., 1999; Kidd et al., 2005) as well as non-spatial cues such as voice characteristics and prosody (e.g., Darwin and Hukin, 2000; Brungart et al., 2001; Darwin, 2008). These features drive selective attention and allow listeners

to focus on one talker of interest while filtering out competing talkers and noise (Shinn-Cunningham, 2008; Carlile, 2014 for reviews).

Cocktail party environments however are rarely static. In a multi-talker exchange, the focus of a conversation constantly shifts from one talker to another. Listeners must be able to rapidly reorient their selective attention in order to follow a conversation. Although non-spatial cues are important in initial auditory grouping, differences in *spatial* location drive temporal streaming particularly in complex multi-talker settings (Shinn-Cunningham, 2005; Allen et al., 2008; Ihlefeld and Shinn-Cunningham, 2008). Little is known about the perceptual consequences of switching *spatial* attention especially in dynamic conversations which involves integration of information across space and time (Sacks et al., 1974; Hutchby and Wooffitt, 2008).

Spatial attention operates like a searchlight, where processing resources can be allocated to a particular region or item in space. This spotlight of attention is limited and there is a gradient where attention falls off as a function of distance from the attended source (Mondor and Zatorre, 1995; Allen et al., 2009). There are benefits of knowing where and when to listen (Kidd et al., 2005; Kitterick et al., 2010) and any deviations from expectancy can lead to a reduction in speech intelligibility. This is consistent with Brungart and Simpson (2007), who showed that performance in a dynamic listening task decreased as a function of spatial transition probability. There has been strong evidence to suggest that auditory attention is object based and that representations build up over time (e.g., Best et al., 2008). Consequently shifts in stimulus location or a change in the attended-to voice result in a cost in streaming performance (Best et al., 2008, 2010).

These studies all highlight the benefit of spatial continuity in auditory object formation and establish that there is a cost associated with switching attention, even when switches are cued and predictable (Best et al., 2010; Koch et al., 2011). Reorientation of spatial attention is critical in the context of following conversations, yet little is known about the processes which drive this. Previous multi-talker studies have been limited to non-complex stimuli such as tones, digits, and simple speech corpora such as the co-ordinate response measure (CRM). However, this is not truly reflective of the cognitive demands of real world listening, which requires multiple element retention and semantic integration across space and time.

Over the past decade, increasing literature has been devoted to unraveling the role of cognition in cocktail party listening (Akeroyd, 2008; Arlinger et al., 2009 for reviews). In particular working memory, the capacity to hold and manipulate task relevant information (Baddeley, 2003; Engle and Kane, 2004), has been central to understanding how we interact with the world around us. Working memory is important for selective attention (de Fockert, 2013), hypothesis generation (Francis and Nusbaum, 2009) and suppressing the effect of distracters (Sörqvist, 2010; Hughes et al., 2013).

Studies in the visual domain (Kane and Engle, 2003; Caparos and Linnell, 2010; Ahmed and de Fockert, 2012) and the auditory domain (Conway et al., 2001; Dalton et al., 2009) have shown that working memory and processing load affect the spatial window of attention. Maintaining task relevant information is dependent on the precision of this selective attention, which influences the degree of distracter processing (Lavie et al., 2004; Lavie, 2005; de Fockert, 2013). As working memory demands increase, performance begins to decline in selective attention tasks, which results in a rise in subjective listening effort (Rönnberg et al., 2014). The recently, proposed ease of language understanding (Rönnberg et al., 2008, 2013) and cognitive spare capacity (Rudner et al., 2011; Mishra et al., 2014) models posit that listeners have a finite pool of working memory resources, which can be allocated to encoding, rehearsal, and comprehension of stimuli. The greater the cognitive load, the less residual resources available for processing of information. Ultimately, complex auditory scenes not only present a challenge in terms of selective attention but also cognitive demands, which influence the fidelity of recall.

This study aimed to investigate the cost of switching spatial attention during conversational turn-taking. We aimed to explore the relationship between attention switching and cognitive processes including working memory in normally-hearing listeners. Word recall and discourse comprehension were examined using matrix sentences (Hagerman, 1982) in a novel paradigm involving speech rehearsal and spatial reorientation. Matrix sentences, which are syntactically fixed but semantically unpredictable, have low stimulus redundancy and allow for the examination of recall independent of context. These structured sentences are particularly appealing for this study as they better approximate the content, semantic diversity and working memory demands of a real world conversation compared to digit recall or predictable closed set sentences found in CRM speech.

Experiment 1 investigated six word recall following a single endogenous switch in spatial attention. Matrix sentences from three concurrent sources were used to isolate spatial switching costs. All three sources were drawn from recordings of the same talker to control for non-spatial cues such as voice characteristics, thereby forcing listeners to rely on spatial information to separate and drive selective attention. Performance in trials involving a switch in target location between two sentences was compared to trials with a non-shifting target. The target location was varied to investigate whether recall differed as a function of the size, spatial hemisphere (left vs. right), and direction of the shift. It was hypothesized that there would be a decrease in recall following a shift in target location, due to a disruption in auditory streaming (Best et al., 2008, 2010) and attentional reorientation following target search (Kidd et al., 2005; Brungart and Simpson, 2007). In addition, cognitive functions including working memory capacity were hypothesized to be correlated with total words recalled and distractor processing during conversational turn-taking.

Experiment 2 was designed as a follow-up to Experiment 1, and investigated the effect of increasing processing load on sentence comprehension. Comprehension of speech relies not only on effective recall but a combination of processes including; segregation of competing streams, discrimination of words, and semantic processing at the sentence level. These processes are important in adverse listening conditions, particularly when listening in demanding situations with high levels of informational masking. This experiment aimed to investigate whether switching performance was load dependent, consistent with a working memory hypothesis. Rather than assessing simple word recall,

this experiment used performance on questions related to the content of the sentences to assess the extent of semantic processing. If working memory is involved in attention switching then we would anticipate an increase in switching cost with increasing question difficulty.

Materials and Methods

Participants

Sixteen young normally-hearing listeners (9 male, aged 21–35, $M = 23.9$, $SD = 4.0$) participated in two auditory attention switching experiments. All listeners had English as their first language, normal hearing as assessed by a pure-tone audiogram (<20 dB hearing loss at frequencies between 250 and 8000 Hz), and no reported cognitive or attentional deficits. All subjects gave written informed consent in accordance with the Human Research Ethics Committee, University of Sydney.

Setup

Experiments were conducted in a sound attenuated audiometric booth (2.5 × 2.4 × 2.2 m in dimension). Listeners sat with their head fixed on a chin rest facing an array of three Tannoy Active loudspeakers, positioned at eye level 1 m from the head at –30, 0, and 30° azimuth.

Stimuli

Three pairs of matrix sentences, recorded from a single male Australian English talker, were presented from the three loudspeaker locations (**Figure 1**). Matrix sentences were syntactically fixed and comprised of name, verb, number, adjective, and noun elements. Sentences were constructed at each trial by randomly sampling each element without replacement from a list of 10 possible words. All words within a trial occurred only once, with the exception of the target name which occurred twice.

Words were 500 ms in duration with the exception of nouns, which were time stretched to 600 ms using Adobe Audition 3.0. This manipulation was applied to reproduce the natural prosodic lengthening of speech at phrase boundaries (Wightman et al., 1992). A 350 ms silence gap was introduced between sentence pairs to replicate the average conversational turn-taking duration of English speech (Stivers et al., 2009). In addition, sentences were staggered with a 50 ms offset to (i) reduce the effects of energetic masking encountered with synchronized concurrent talkers and (ii) enhance grouping by staggering onsets. Offset combinations were randomized each trial and balanced for all locations. Stimuli were generated using Matlab (MathWorks) and played through an RME FireFace UCX soundcard at 48 kHz sampling rate. All sentences were presented at 65 dB SPL.

Procedure

Both experiments utilized the same setup and stimuli but differed in their post stimulus task. Each trial began with a 0.75 s 500 Hz priming tone presented from one of three loudspeakers. Subjects directed their spatial attention to this cue and were instructed to remember the name and sentence that followed at this location. A second set of matrix sentences were presented after a silent turn-taking gap. Subjects were required to search for and attend to the sentence with the same target name, which either remained in the same spatial location (no switch trials) or moved to another spatial location (switch trials). There were three possible target locations for the first sentence (S1) and three possible target locations for the second sentence (S2), yielding a total of nine possible

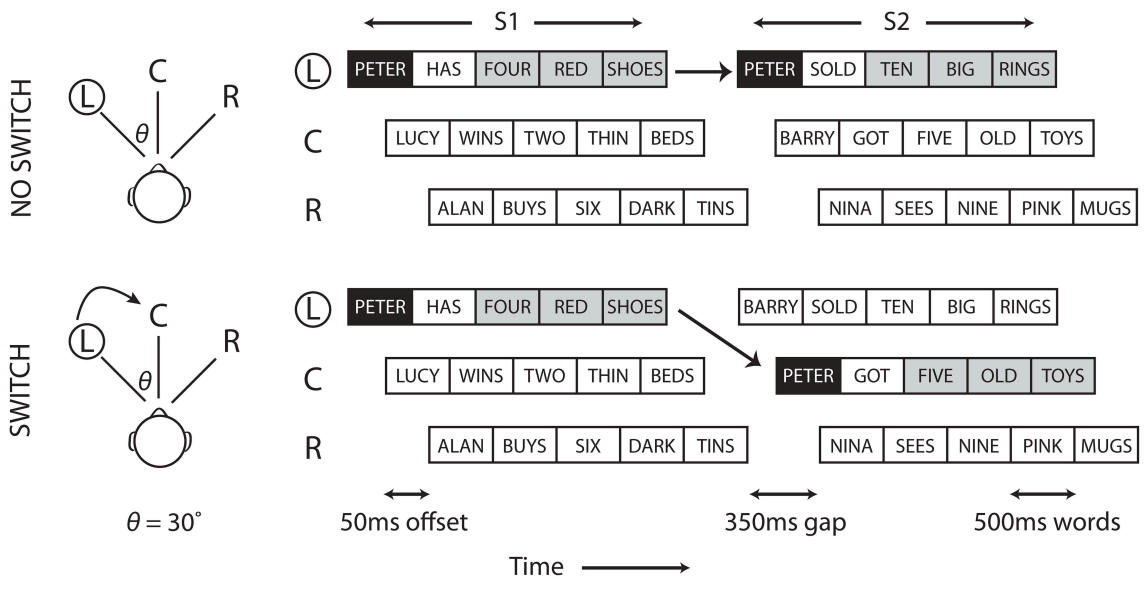


FIGURE 1 | Experimental setup. Matrix sentences were presented over loudspeakers positioned to the left (L), center (C), and right (R) of the listener's head. Examples are shown for a single no switch (top), and switch (bottom) trial. Subjects attended to the cued location (circled) and followed sentences (S1 and S2) with the cued name, in this case "Peter". In Experiment 1, subjects were required to verbally recall the last three words of each target sentence (gray). In Experiment 2, subjects answered a graded multiple choice question related to the target sentences.

spatial conditions. The target sentence was presented with equal likelihood at all loudspeaker locations. Subjects performed one of two tasks at the end of each trial depending on the experiment.

Experiment 1: Speech Recall

Subjects were required to verbally report the last three words of each target sentence in correct serial order (six item recall). Subjects also reported the target name to verify that they followed the correct stream. Only trials where subjects correctly identified the name were included in analysis (83.5% of trials). Verbal responses were recorded using a microphone and saved for scoring and analysis after the experiment. If a subject could not recall a word during a trial, this was registered as a "pass". Subjects completed a short training block to familiarize themselves with the stimuli and procedure before starting a total of 24 repeats for nine spatial conditions in randomized order (4 blocks of 54 trials).

Experiment 2: Speech Comprehension

Subjects were presented with two multiple choice questions on a computer screen following each trial (one for each target sentence). Questions varied in complexity ranging from 1-Step simple recognition questions e.g., which word was in the target sentence? to 2-Step specific recall questions e.g., which big item did Peter sell? to 3-Step quantity comparison questions e.g., which item had the smallest/largest number?. These questions were based on those used by Rönnberg et al. (2014). Subjects were required to respond as fast and as accurately as possible using a keypad. Subjects participated in a short training block before completing total of 6 repeats for 9 spatial conditions and 3 question types in randomized order (3 blocks of 54 trials).

Cognitive Tasks

Subjects also completed a battery of cognitive tests including the reading span test to measure working memory capacity (Daneman and Carpenter, 1980; Baddeley et al., 1985), attention network test to measure attentional modulation (Fan et al., 2002), and trail making test to measure executive function (Reitan, 1958).

In the reading span test, subjects were presented with a series of short sentences on a computer screen, starting with 3 and increasing in length to 6. Participants were required to read the sentences out aloud and verbally report whether each made literal sense or not (half were non-sensical). At the end of each series, subjects were prompted to recall either the first or last words of each of the sentences. The number of total correct words recalled was used as a measure of working memory capacity.

The attention network test was a cued reaction time flanker task presented on a computer screen. Subjects attended to a fixation cross at the center of the screen which was accompanied by an arrow above or below the fixation point. Subjects were required to respond as fast and as accurately with the left or right keyboard keys to indicate the direction of the arrow. A number of conditions were tested including with congruent/incongruent flanking arrows, with/without a temporal alerting cue, and with/without a target spatial cue. Three measures of attentional control were extracted from the test; alerting ability, orienting ability, and cue conflict resolution.

The trail making test consisted of two timed pen and paper tests which required subjects to connect a series of labeled circles in ascending numerical order (Trail-A) or alternating numeric-alphabetical order (Trail-B). These two tests provide coarse measures of visuo-motor processing and task switching speed, respectively, while the difference score (Trail-B minus Trail-A) provides an estimate of executive control ability (Sánchez-Cubillo et al., 2009).

Data Analysis

Center Correction

A score correction was applied to all conditions containing a central target, to account for the energetic disadvantage posed by the absence of an acoustic "better-ear" (Zurek, 1993). This disadvantage was estimated for each subject as the difference between the central no switch condition (CC) and the mean of the left (LL) and right (RR) no switch conditions. The full correction was applied to the CC condition, while half of this correction was applied to conditions which contained one central target (LC, CL, CR, and RC).

Error Analysis

In addition to measuring the number of words correct, the errors committed by each subject were analyzed for their relative frequency. In Experiment 1, "masker confusions" and "passes" were calculated for each condition to quantify the degree of informational masking and failures in word recall, respectively. Masker confusions were instances where a subject reported a word presented in a concurrent masking stream, while passes were instances where a subject failed to register a response for a particular word.

In Experiment 2, subjects were presented with multiple choice questions with one correct option and two incorrect options. For 1-step questions, incorrect options included a masker confusion and an unspoken word (a word which was not presented in the trial and was reflective of random guessing). For 2- and 3-step questions, incorrect options included a masker confusion and a sentence order confusion (a word which was present in the target stream but was embedded in the alternate sentence). The latter type of error occurred when subjects mixed words from sentence 1 and 2, reflecting a failure to integrate information.

Statistical Analysis

Data from Experiment 1 were normally distributed. The mean number of words correct for each spatial condition were compared using a repeated measures One-Way ANOVA. No switch trials were compared with corresponding switch trials using a series of planned paired *t*-tests. Switching costs were calculated for each subject as the mean difference in performance between no switch and switch conditions. Further analysis was performed on recall rates, masker confusions, and passes using Three-Way repeated measures ANOVAs examining the effects of word, sentence position, and switching. The relationship between listening task performance and cognitive test scores were examined using linear correlations.

Data from Experiment 2 were not normally distributed and were arc-sine transformed. This transformation converts binomial data into an approximately normal distribution for parametric analysis (Studebaker, 1985). Performance was analyzed using Two-Way repeated measures ANOVAs with task difficulty and switching as independent variables. The difference between switching and no switching performance were analyzed for each question type using paired *t*-tests. Outliers were not removed from either experiment.

Results

Experiment 1

Total Words Recalled

There was considerable variability in performance between individuals in the speech recall task (**Figure 2A**). Scores ranged from 1.7 to 5.8 words correct per trial (out of 6), with differences as large as twofold between subjects in certain conditions. Despite this variability, trends across conditions were similar, with mean performance higher in no switch trials compared to switch trials.

Scores were consistently higher for some subjects than for others. To better examine the within-subjects effect of switching, the number of words correct was normalized to the maximum score for each subject (**Figure 2B**). A One-Way repeated measures ANOVA on normalized data confirmed a significant effect of spatial condition [$F_{(4.5,\ 67.9)} = 12.5$, $p < 0.001$]. Planned pairwise comparisons indicated a significant recall advantage in no switch trials compared to respective switch trials (LL > LC, LR; CC > CL, CR; RR > RL, RC). There were no significant differences between any of the switch conditions, demonstrating no effect of location, direction, and angular size of the spatial shift on word recall. Overall, switching spatial attention resulted in a 10.7 ± 1.3% decrease in word recall when averaged across subjects and locations.

Sentence and Word Recall

A Three-Way repeated measures ANOVA on percent correct data revealed a significant main effect of sentence number [$F_{(1,\ 15)} = 20.0$, $p < 0.001$], word position [$F_{(2,\ 30)} = 6.3$, $p < 0.01$], and switching [$F_{(1,\ 15)} = 69.0$, $p < 0.001$]. Recall was lower for the second target word and for the second target sentence (S2) in each trial, particularly following a switch in spatial attention (**Figure 3A**). There was a significant sentence by switch interaction effect [$F_{(1,\ 15)} = 10.8$, $p < 0.01$], where recall dropped significantly between S1 (71.6 ± 4.5%) and S2 (51.9 ± 4.6%) in the switch condition ($p < 0.001$). In contrast, there was minimal decline in recall between S1 and S2 in the no switch condition (76.4 ± 4.0 vs. 68.4 ± 3.4%, respectively, $p > 0.05$).

The effect of word position resembled a classic serial position curve (**Figure 3A**), with recall greatest for the first and last items in each sentence. A significant sentence by word interaction effect [$F_{(1,\ 30)} = 3.8$, $p < 0.05$] was observed, where the final target word was recalled significantly more often than the second target word (68.6 ± 2.3 vs. 60.5 ± 3.8%, $p < 0.01$) for S2 only. This word recency effect was less pronounced in S1 but was observed in both switch and no switch conditions.

Masker Confusions

A Three-Way repeated measures ANOVA on masker confusions revealed a significant main effect of sentence number [$F_{(1,\ 15)} = 36.9$, $p < 0.001$], word position [$F_{(2,\ 30)} = 21.0$, $p < 0.001$], and switching [$F_{(1,15)} = 12.3$, $p < 0.01$], and a significant sentence by switch interaction effect [$F_{(1,15)} = 17.9$, $p < 0.01$]. Masker confusions constituted ~9–19% of responses and were most prevalent in the final word of each sentence, and primarily in S2 (**Figure 3B**). There was no significant difference in the frequency of masker confusions between sentences in the no switch condition. However, the number of masker confusions doubled from S1 (9.2 ± 1.2%) to S2 (18.8 ± 1.8%) following a switch in spatial attention ($p < 0.001$).

Interestingly, listeners demonstrated significantly greater masker confusions for the last word (15.1 ± 1.6%), compared to the first (10.9 ± 1.2%, $p < 0.001$) and the second target words (12.2 ± 1.1%, $p < 0.01$) in each sentence. We speculate that this may be a "masker" recency effect.

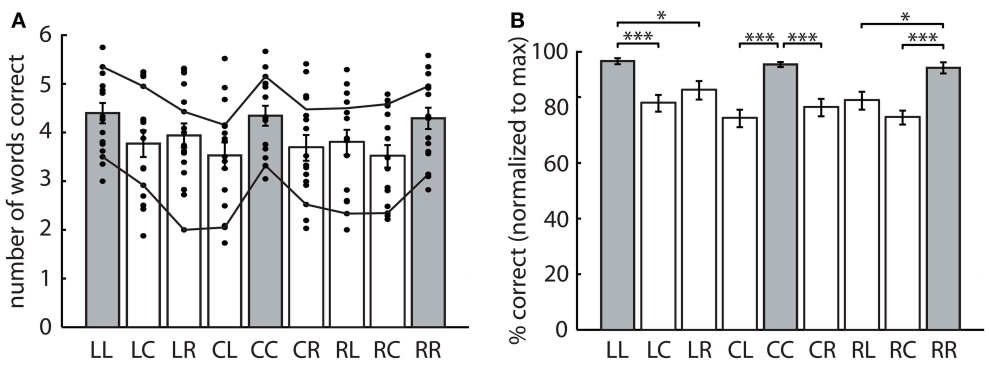


FIGURE 2 | Experiment 1 scores. (A) Mean number of words correct for nine spatial conditions. Large variability in individuals scores (dots) was observed but trends across conditions were similar between high and low performing subjects (solid lines). **(B)** Normalized percentage correct for nine spatial conditions. Performance was significantly higher for no switch (gray) compared to switch trials (white). Bars represent mean ± SEM, $^{*}p < 0.05$, $^{***}p < 0.001$.

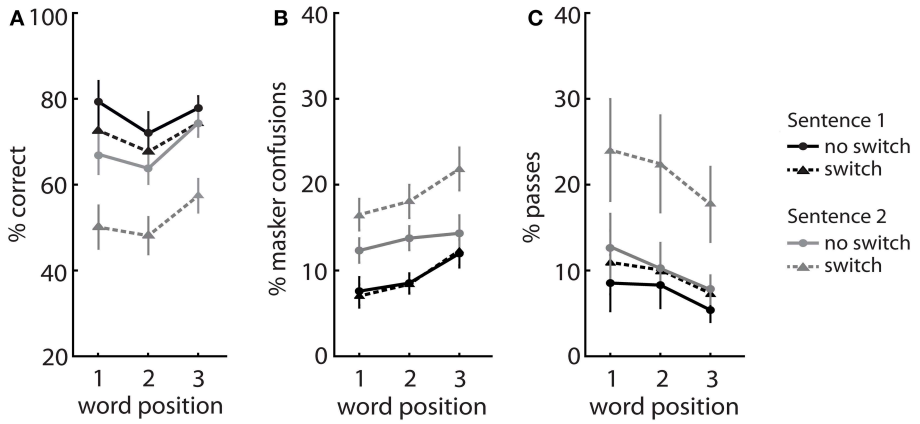


FIGURE 3 | Experiment 1—The effect of sentence and word position on (A) words correct, (B) masker confusions, and (C) frequency of passes, during no switch and switch trials. Data presented as mean ± SEM.

TABLE 1 | Pearson correlation coefficients between Experiment 1 scores and cognitive test scores.

	RST score	ANT-A	ANT-O	ANT-C	Trail-A	Trail-B	Trail B-A
Words correct	**0.457***	0.057	0.219	−0.320	−0.390	**-0.458***	−0.307
Switching cost	**-0.442***	−0.316	−0.252	0.176	0.389	0.396	0.245
Masker confusions	0.423	0.054	0.109	−0.043	−0.019	0.348	0.361
Passes	**-0.565***	−0.261	−0.290	0.417	0.413	0.166	0.002

*RST, Reading Span Test; ANT, Attention Network Test Alerting (ANT-A); Orienting, (ANT-O); Conflict resolving ability, (ANT-C); Trail, Trail making test A (Trail-A), test B (Trail-B), and difference score (Trail B-A). *p < 0.05 shown in bold.*

Passes

A Three-Way repeated measures ANOVA on pass rates revealed a significant main effect of sentence number [$F_{(1, 15)} = 10.6$, $p < 0.01$], word position [$F_{(1.3, 18.9)} = 4.2$, $p < 0.05$], and switching [$F_{(1, 15)} = 14.1$, $p < 0.01$], and a significant sentence by switch interaction effect [$F_{(1, 15)} = 7.2$, $p < 0.05$]. Passes were more prevalent for the first word of each sentence, and for S2 overall (**Figure 3C**). The frequency of passing remained below 12% in the no switch condition, and there was no significant difference in pass rates between the first and second sentences (7.4 ± 2.4 vs. $10.3 \pm 2.4\%$, $p > 0.05$). However, the likelihood of passing increased twofold for S2 when there was a switch ($21.4 \pm 5.2\%$), compared to S1 pre-switch ($9.4 \pm 3.0\%$, $p < 0.05$).

Passes in the second sentence were not always due to a failure in search. Subjects were able to recall at least one correct word from S2 in 87.7% of no switch trials and 69.5% of switch trials. This implies that they were able to locate the second sentence in the majority of trials. A supplementary experiment was devised using the same paradigm but without recall of elements, to test the ability to simply follow the target with minimal cognitive load. In this experiment, a subset of six subjects was able to locate S2 with a high success rate, 93.1% of the time during no switch trials and 88.4% of the time during switch trials.

Cognitive Correlates

Correlations between Experiment 1 performance and cognitive test scores for the cohort are shown in **Table 1**. The number of words correct per trial were positively correlated with reading span score ($r = 0.46$, $p < 0.05$) and negatively correlated with Trail-B time ($r = -0.46$, $p < 0.05$). Reading span score was also negatively correlated with switching costs ($r = -0.44$, $p < 0.05$) and frequency of passes ($r = -0.57$, $p < 0.05$). There were no significant correlations between any measure of the attention network test and performance in the listening task. Other measures of the trail making test were also not correlated with listening performance.

Experiment 2

Percent Correct

Experiment 2 was designed as a follow-up to Experiment 1, to explore the effect of increasing processing load on switching costs. A Two-Way repeated measure ANOVA on percent correct data revealed a significant main effect of switching [$F_{(1, 15)} = 21.2$, $p < 0.001$], and a main effect of question type [$F_{(2, 30)} = 12.7$, $p < 0.001$] on correct responses, but no significant question by switch interaction. Sentence comprehension decreased for switch trials and decreased with increasing question complexity (**Figure 4**). Switching costs were load dependent and increased proportionally with the number of cognitive operations in each question (6.9, 8.5, and 9.2% cost for 1-step, 2-step, and 3-step questions, respectively). Planned pairwise comparisons revealed a significant switching cost only in the 2-step [$t_{(15)} = 3.4$, $p < 0.05$] and 3-step conditions [$t_{(15)} = 2.2$, $p < 0.05$] but not in the 1-step condition [$t_{(15)} = 2.3$, $p > 0.05$].

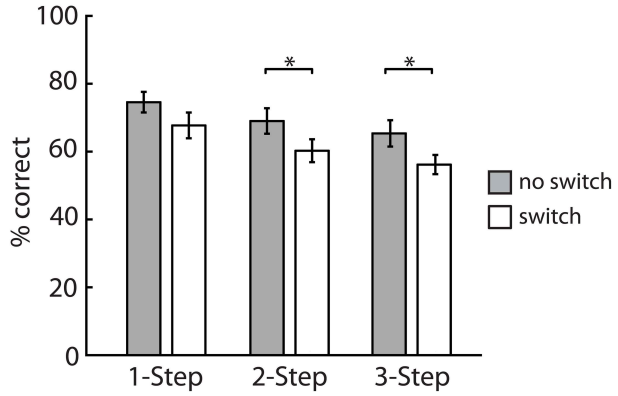


FIGURE 4 | Experiment 2—Percentage of correct responses for three question types. Switching costs were significant only for 2 and 3-step questions. Bars represent mean ± SEM, *$p < 0.05$.

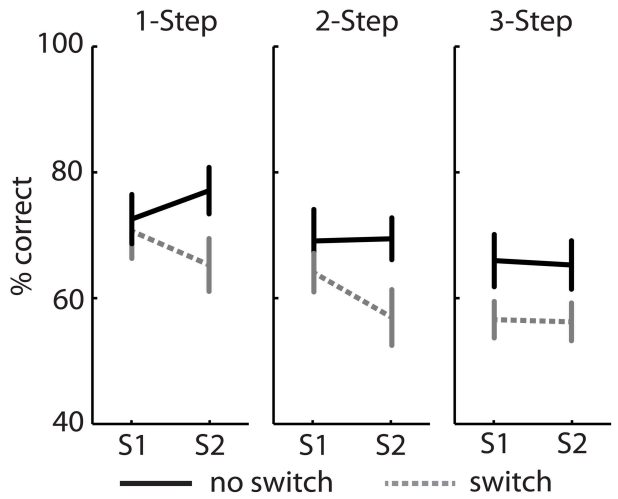


FIGURE 5 | Experiment 2—Percentage of correct responses from sentence 1 (S1) and 2 (S2) for three question types. Bars represent mean ± SEM.

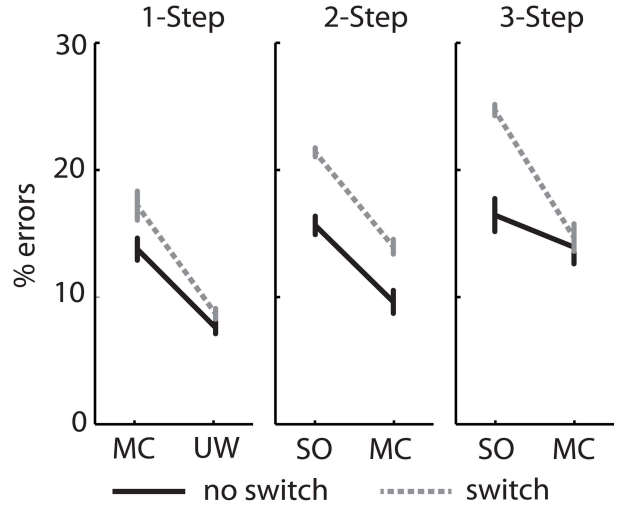


FIGURE 6 | Experiment 2—Proportion of masker confusion (MC), unspoken word (UW), and sentence order (SO) errors for three question types. Data presented as mean ± SEM.

Sentence Analysis

A Three-Way repeated measures ANOVA revealed that there was no significant main effect of sentence on performance [$F_{(1,15)} = 1.1$, $p = 0.3$]. The sentence by switch interaction was non-significant [$F_{(1,15)} = 3.1$, $p = 0.098$]. As seen in **Figure 5**, performance was higher for S1 compared to S2 only under certain conditions. Trends were similar to those observed in Experiment 1 with a small sentence primacy effect evident following a switch in both 1-step and 2-step conditions. This effect was however abolished following a complex 3-step question.

Error Analysis

There were greater errors committed in the switch condition compared to the no switch condition (**Figure 6**). For simple 1-step questions, subjects were more likely to report masker confusions than unspoken words (with a guess rate of <10%). For complex 2- and 3-step questions, sentence order confusions were more prevalent than masker confusions. Switching spatial attention increased the proportion of all error types in the 1- and 2-step conditions. However, in the 3-step condition, switching resulted in a disproportionate increase in sentence order confusions but not of masker confusions. Thus, as question load increased, subjects tended to make less location attribution errors (confusing competing streams) and more semantic attribution errors (confusing elements from S1 to S2).

Discussion

This study examined the cost of switching endogenous spatial attention in a dynamic three talker cocktail party setting. In a cohort of young normally-hearing listeners, there was a significant decrease in word recall and discourse comprehension following a switch in target location in a two sentence selective attention paradigm. The cost was independent of the location, direction, and angular size of the spatial shift and was predominantly confined to the second sentence post switch. The drop in recall was associated with a concomitant increase in reported masker confusions and word omissions. The significant relationship between listening task performance and reading span score supports the hypothesis that switching efficacy is driven by working memory. An individual's working memory capacity impacts their ability to accurately recall words across space and time. This study also demonstrates that there is a cognitive load associated with switching attention during conversational listening. Systematic increases in question difficulty lead to a progressive decline in switch performance, providing evidence that attention switching is both load and working memory dependent.

The Cost of Switching Spatial Attention

Switching spatial attention resulted in a decrease in word recall. The costs observed in this study are within range of previous

reported switching costs of 5–15% by Best et al. (2008) using five talkers, and up to 15% observed by Brungart and Simpson (2007) using three talkers. One key difference in this study is the use of a single male talker at all three locations, to control for the influence of non-spatial cues including voice characteristics. Although not ecological, this manipulation isolates the cost of a single endogenous switch in spatial attention using relatively diverse conversational stimuli.

Previous studies have attributed switching costs to target location uncertainty (Kidd et al., 2005; Brungart and Simpson, 2007) and disruption to object streaming continuity (Best et al., 2008, 2010). The reduction in recall, predominantly confined to the second sentence, post switch, supports this notion. This drop in performance however, cannot be solely attributed to a failure to re-engage or find the second sentence as subjects could report at least one correct word from this sentence 78.6% of the time. This implies that cost in this paradigm was not primarily due to location uncertainty, but perhaps other factors such as disruption to streaming or cognitive load. Indeed in a supplementary experiment, without any cognitive load, subjects were able to localize the second target sentence with 90.7% accuracy.

We propose two possible mechanisms for this degradation in second sentence recall. Firstly, cognitive load from word rehearsal may decrease efficiency of the switch and subsequent search for S2. Based on the difference between no load and load identification of S2 (88.4 vs. 69.5%), there does appear to be some evidence for a degradation in localization as a result of rehearsal. Consequently, subjects were more likely to commit masker confusions or pass in S2 as they were unable to identify the target stream. Alternatively, the reduction in recall fidelity may be due to increased cognitive load induced by the switch itself. In Experiment 2, we see that switch costs are not uniform and are load dependent. Systematic increases in post presentation question difficulty amplified the cost of switching, supporting a limited working memory model. Furthermore, analysis of the errors revealed the prevalence of sentence order confusions over masker confusions implying successful stream segregation but unsuccessful attribution of semantic details. Thus, subjects were able to localize the correct target during the switch but were unable to integrate information in the post-stimulus decision phase. This is strong support for the notion of switching increasing cognitive load.

Neither the distance, direction, nor location of the spatial shift had any significant bearing on performance. These results are consistent with the findings of Mondor and Zatorre (1995) and Brungart and Simpson (2007), who demonstrated that performance in a spatial orienting task did not decline as a function of shift distance and angular displacement size. It appears that the average turn taking gap of 350 ms (Stivers et al., 2009) was sufficient to allow subjects to reorient their attention up to 60° in this paradigm. This duration is well outside the timeframe of 80–200 ms proposed by other studies for spatial reorientation (Teder-Salejarvi and Hillyard, 1998). Under the no-load condition, subjects were able to redirect their attention with high success rate during the conversational gap. It was only under load that this performance decreased. Time is a critical factor for speech understanding (Singh et al., 2008, 2013; Koch et al., 2011; Dhamani et al., 2013), particularly in multi-talker conversations which involve rapid and unpredictable shifts in target location.

In Experiment 1, the lack of an interaction between word position score and switching condition demonstrates that there was no temporal impact of the switch on word recognition immediately post switch. Koch et al. (2011) showed that there was a delay associated with having to switch attention between ears in a dichotic listening task. However, there was no significant "inertia" observed in our performance data. The uniform drop in recall across all three words suggests that elements were equally susceptible to interference rather than a failure to reorient attention fast enough.

Even though we did not observe any location dependent costs, it should be noted that scores observed in this study were adjusted with a center correction. The center correction is an estimate of the energetic disadvantage posed by the absence of a better ear. This correction may however *overestimate* the performance disadvantage posed by a central talker flanked by two maskers, and thus *underestimate* the true switching cost when presented with a central target. In addition, the performance disadvantage may not be additive in all switching conditions.

Individual Differences

Notably, we found large individual differences in task performance in this cohort of young normally-hearing listeners. Correlations between switching performance and individual cognitive measures strongly support the theory that working memory is important for maintaining task relevant information in adverse listening conditions (Baddeley, 2003; Engle and Kane, 2004). The positive correlation between number of words correct and working memory capacity reinforces the importance of information retention and manipulation for comprehension during dynamic conversations. Furthermore, the negative correlation between switching costs and working memory highlights the disparity between high and low working memory individuals in their ability to retain information across switches. High working memory subjects are not only better at selective attention tasks (Conway et al., 2001) but have been shown to be more proficient at divided attention tasks which involve monitoring the occurrence of a target name across multiple streams (Colflesh and Conway, 2007).

However, contrary to previous predictions, working memory was not associated with distractor processing as suggested by some studies (Conway et al., 2001; Ahmed and de Fockert, 2012). Switching attention did increase the overall proportion of masker confusions (**Figure 3B**), but this was not associated with individual cognitive correlates. This may be due to the type of distraction encountered in this task. Recent studies propose a duplex theory of distraction which posit that an irrelevant stream can either (i) capture attention due to stimulus deviation or (ii) interfere with serial rehearsal due to the changing state of the distractor stream (Hughes, 2014; Sörqvist and Rönnberg, 2014). The former, but not the latter, has been shown to be correlated with working memory capacity (Sörqvist et al., 2013). It is quite possible that non-target streams interfered with the process of rehearsal rather than attention capture in this experiment. Another potential explanation may lie in the nature

of the task, which permitted the absence of responses ("passes"). Reporting masker confusions was thus dependent on the discretion and response criterion of the subject. Interestingly, analysis of pass frequency was associated with the second sentence switching performance (**Figure 3C**) and negatively correlated with an individual's working memory (**Table 1**). This suggests that decay of rehearsed information found in this study may be related to information storage capacity.

The other significant correlation was between total words recalled and Trail-B time, which is a measure of task switching ability (Sánchez-Cubillo et al., 2009). Perhaps not surprisingly, faster task switching meant better performance in our listening task- which inherently involves a switch from selective to divided, back to selective attention. Visuo-motor processing (Trail-A) and executive function (Trail B–A) were perhaps not as prominent in this listening task, however some correlations were bordering significance.

The lack of a correlation between any of the measures of the attention network test may have two explanations. While the ANT may be effective in revealing differences in clinical populations such as in ADHD (Johnson et al., 2008), the test has less resolution in this cohort of young healthy subjects. Secondly, the test examines basic attentional modulation and not attentional capacity under load, the latter of which is most important when dealing with multi-talker cocktail party environments. Neither of the three measures of the ANT were driving the effects we were observing in our listening test, which were primarily working memory and task switching based. It should also be noted that the tests employed in this study are not mutually exclusive and there may be some overlap between cognitive processes.

Furthermore, differences in performance may depend on the type of strategy adopted by the individual listener. Studies have shown that the probability of target locations has an influence on the allocation of attention and consequently speech intelligibility (Kidd et al., 2005; Brungart and Simpson, 2007). The current task, where all locations are equally probable as targets, requires both selective and divided attention, and is reflective of an unpredictable, uncued conversation. Interestingly, subjects were found to distribute their expectations evenly during the conversational gap. Following a switch in target location, almost half (48.9%) of reported masker confusions in S2 arose from the original S1 target location while the other 51.1% originated from the non-target location. This provides evidence that the no switch advantage was not due to subjects simply keeping their attention fixated on the S1 location.

The Importance of Working Memory

Working memory involves the storage, manipulation and recall of goal-relevant information, and the inhibition of distracters. This study reinforces the notion of conversational tracking as an active task which requires cognitive resources, especially when there is a shift in spatial attention. This supports both the cognitive spare capacity model proposed by Rudner et al. (2011) and ease of language understanding model by Rönnberg et al. (2008, 2013).

Based on these models, working memory is limited and must be allocated to various components of the listening task. Here working memory is important in encoding, rehearsing, and recalling information across switches in spatial attention. Individuals with low working memory capacity can only encode a limited amount of information and have little residual "spare capacity" to process the information, hence lower recall. The introduction of a switch requires allocation of cognitive resources and further limits spare capacity to encode and recall information particularly in S2. Individuals with high working memory capacity experience these constraints to a lesser extent. Furthermore, studies have shown that subjects with better cognitive abilities including working memory, distracter inhibition, and text reception threshold have better speech intelligibility, selective attention, and word recall in noise (Kjellberg et al., 2008; Koelewijn et al., 2012; Meister et al., 2013).

In Experiment 2, increases in cognitive load had implications for broader discourse comprehension. Based on the ease of language understanding model, higher working memory load leads to a decrease in the fidelity of encoded information which impacts lexical access and downstream comprehension (Rönnberg et al., 2008, 2013). This has implications not only for normally-hearing listeners but for elderly and hearing impaired listeners with peripheral and cognitive deficits. Working memory deteriorates with age and there is greater cognitive load and effort following hearing loss (Tun et al., 2009). Peripheral deficits lead to a myriad of downstream deficits including elevated thresholds, failure to group and segregate sounds, poorer speech intelligibility, and greater central processing demands.

In real world listening we rely on semantic information and contextual cues to endogenously guide attention (Pichora-Fuller et al., 1995; Meister et al., 2013). The use of a fixed syntax, unpredictable corpus allows for examination of sentence comprehension while removing the influence of context. While this is advantageous in a controlled environment for isolating recall costs, in real world situations context plays an important role in stream formation and discourse comprehension (Pichora-Fuller et al., 1995). Context is believed to alleviate some of this cognitive load associated with listening in adverse conditions as it allows for top-down prediction of words (Rönnberg et al., 2013).

Another potential contributing factor not measured in this experiment is the level of proactive interference experienced by each subject. Proactive interference refers to the degradation of memory traces by prior encoded information (Kane and Engle, 2000), particularly items with a similar context—such as, words within the same category in a closed set corpus. The ability to resist semantic proactive interference has been shown to be closely related to speech in noise recognition (Ellis and Rönnberg, 2014). Differences in this study in the level of proactive interference between high and low working memory participants may mediate cross-trial or within-trial interference and hence the accuracy of recall. The increase in masker confusions and sentence order confusions following a switch may reflect an increase in interference from previously encoded sentences. However, these errors are difficult to quantify in the current study as the same words can be present in multiple successive trials.

Conclusion

Switching spatial attention in a cocktail party setting imposes a cognitive load which impacts short term recall of words. This cognitive load impacts the disengagement and reorientation of attention and consequently the encoding of information immediately following the switch. This has a downstream effect on comprehension of sentences in a multi-talker conversation. Switching led to an increase in distractor interference and higher likelihood to miss words. Costs appear to be direction, spatial hemisphere, and size independent but do seem to be load dependent and only significant with tasks involving multiple operations. These results support the notion of a limited working memory model which is involved in directing spatial attention, encoding, and post-perceptual processing of stimuli in a multi-talker auditory scene.

Acknowledgments

The authors gratefully acknowledge Virginia Best for her assistance in the development of the experimental paradigm and Thomas Lunner for his helpful feedback.

References

Ahmed, L., and de Fockert, J. W. (2012). Focusing on attention: the effects of working memory capacity and load on selective attention. *PLoS ONE* 7:e43101. doi: 10.1371/journal.pone.0043101

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

Allen, K., Alais, D., and Carlile, S. (2009). Speech intelligibility reduces over distance from an attended location: evidence for an auditory spatial gradient of attention. *Percept. Psychophys.* 71, 164–173. doi: 10.3758/APP.71.1.164

Allen, K., Carlile, S., and Alais, D. (2008). Contributions of talker characteristics and spatial location to auditory streaming. *J. Acoust. Soc. Am.* 123, 1562–1570. doi: 10.1121/1.2831774

Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201

Baddeley, A., Logie, R., Nimmo-Smith, I., and Brereton, N. (1985). Components of fluent reading. *J. Mem. Lang.* 24, 119–131. doi: 10.1016/0749-596X(85)90019-1

Best, V., Ozmeral, E. J., Kopco, N., and Shinn-Cunningham, B. G. (2008). Object continuity enhances selective auditory attention. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13174–13178. doi: 10.1073/pnas.0803718105

Best, V., Shinn-Cunningham, B. G., Ozmeral, E. J., and Kopco, N. (2010). Exploring the benefit of auditory spatial continuity. *J. Acoust. Soc. Am.* 127, EL258–EL264. doi: 10.1121/1.3431093

Brungart, D., and Simpson, B. (2007). Cocktail party listening in a dynamic multitalker environment. *Percept. Psychophys.* 69, 79–91. doi: 10.3758/BF03194455

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J. Acoust. Soc. Am.* 110, 2527. doi: 10.1121/1.1408946

Caparos, S., and Linnell, K. J. (2010). The spatial focus of attention is controlled at perceptual and cognitive levels. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1080–1107. doi: 10.1037/a0020367

Carlile, S. (2014). Active listening: speech intelligibility in noisy environments. *Acoust. Aust.* 42, 90–96.

Colflesh, G., and Conway, A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychon. Bull. Rev.* 14, 699–703. doi: 10.3758/BF03196824

Conway, A. R. A., Cowan, N., and Bunting, M. F. (2001). The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychon. Bull. Rev.* 8, 331–335. doi: 10.3758/BF03196169

Dalton, P., Santangelo, V., and Spence, C. (2009). The role of working memory in auditory selective attention. *Q. J. Exp. Psychol. (Hove).* 62, 2126–2132. doi: 10.1080/17470210903023646

Daneman, M., and Carpenter, P. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 466, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Darwin, C. J. (2008). Listening to speech in the presence of other sounds. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 363, 1011–1021. doi: 10.1098/rstb.2007.2156

Darwin, C., and Hukin, R. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *J. Acoust. Soc. Am.* 107, 970–977. doi: 10.1121/1.428278

de Fockert, J. W. (2013). Beyond perceptual load and dilution: a review of the role of working memory in selective attention. *Front. Psychol.* 4:287. doi: 10.3389/fpsyg.2013.00287

Dhamani, I., Leung, J., Carlile, S., and Sharma, M. (2013). Switch attention to listen. *Sci. Rep.* 3:1297. doi: 10.1038/srep01297

Ellis, R. J., and Rönnberg, J. (2014). Cognition and speech-in-noise recognition: the role of proactive interference. *J. Am. Acad. Audiol.* 25, 975–982. doi: 10.3766/jaaa.25.10.6

Engle, R. W., and Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychol. Learn. Motiv.* 44, 145–199. doi: 10.1016/S0079-7421(03)44005-X

Fan, J., McCandliss, B., and Sommer, T. (2002). Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* 14, 340–347. doi: 10.1162/089892902317361886

Francis, A., and Nusbaum, H. (2009). Effects of intelligibility on working memory demand for speech perception. *Attent. Percept. Psychophys.* 71, 1360–1374. doi: 10.3758/APP.71.6.1360

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *J. Acoust. Soc. Am.* 106, 3578–3588. doi: 10.1121/1.428211

Griffiths, T. D., and Warren, J. D. (2004). What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892. doi: 10.1038/nrn1538

Hagerman, B. (1982). Sentences for testing speech intelligibility in noise. *Scand. Audiol.* 11, 79–87. doi: 10.3109/01050398209076203

Hughes, R. W. (2014). Auditory distraction: a duplex-mechanism account. *PsyCh J.* 3, 30–41. doi: 10.1002/pchj.44

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., and Jones, D. M. (2013). Cognitive control of auditory distraction: impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 539–553. doi: 10.1037/a0029064

Hutchby, I., and Wooffitt, R. (2008). *Conversational Analysis, 2nd Edn.* Cambridge: Polity.

Ihlefeld, A., and Shinn-Cunningham, B. G. (2008). Disentangling the effects of spatial cues on selection and formation of auditory objects. *J. Acoust. Soc. Am.* 124, 2224–2235. doi: 10.1121/1.2973185

Johnson, K. A., Robertson, I. H., Barry, E., Mulligan, A., Dáibhis, A., Daly, M., et al. (2008). Impaired conflict resolution and alerting in children with ADHD: evidence from the Attention Network Task (ANT). *J. Child Psychol. Psychiatry* 49, 1339–1347. doi: 10.1111/j.1469-7610.2008.01936.x

Kane, M., and Engle, R. (2000). Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 336–358. doi: 10.1037/0278-7393.26.2.336

Kane, M. J., and Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to stroop interference. *J. Exp. Psychol. Gen.* 132, 47–70. doi: 10.1037/0096-3445.132.1.47

Kidd, G., Arbogast, T. L., Mason, C. R., and Gallun, F. J. (2005). The advantage of knowing where to listen. *J. Acoust. Soc. Am.* 118, 3804–3815. doi: 10.1121/1.2109187

Kitterick, P. T., Bailey, P. J., and Summerfield, A. Q. (2010). Benefits of knowing who, where, and when in multi-talker listening. *J. Acoust. Soc. Am.* 127, 2498–2508. doi: 10.1121/1.3327507

Kjellberg, A., Ljung, R., and Hallman, D. (2008). Recall of words heard in noise. *Appl. Cogn. Psychol.* 22, 1088–1098. doi: 10.1002/acp.1422

Koch, I., Lawo, V., Fels, J., and Vorländer, M. (2011). Switching in the cocktail party: exploring intentional control of auditory selective attention. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1140–1147. doi: 10.1037/a0022189

Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J., and Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *Int. J. Otolaryngol.* 2012:865731. doi: 10.1155/2012/865731

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends Cogn. Sci.* 9, 75–82. doi: 10.1016/j.tics.2004.12.004

Lavie, N., Hirst, A., de Fockert, J. W., and Viding, E. (2004). Load theory of selective attention and cognitive control. *J. Exp. Psychol. Gen.* 133, 339–354. doi: 10.1037/0096-3445.133.3.339

Meister, H., Schreitmüller, S., and Grugel, L. (2013). Cognitive resources related to speech recognition with a competing talker in young and older listeners. *Neuroscience* 232, 74–82. doi: 10.1016/j.neuroscience.2012.12.006

Mishra, S., Stenfelt, S., Lunner, T., Ronnberg, J., and Rudner, M. (2014). Cognitive spare capacity in older adults with hearing loss. *Front. Aging Neurosci.* 6:96. doi: 10.3389/fnagi.2014.00096

Mondor, T. A., and Zatorre, R. J. (1995). Shifting and focusing auditory spatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 387–409. doi: 10.1037/0096-1523.21.2.387

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–608. doi: 10.1121/1.412282

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Percept. Mot. Skills* 8, 271–276. doi: 10.2466/pms.1958.8.3.271

Rönnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S99–S105. doi: 10.1080/14992020802301167

Rönnberg, N., Rudner, M., Lunner, T., and Stenfelt, S. (2014). Assessing listening effort by measuring short-term memory storage and processing of speech in noise. *Speech Lang. Hear.* 17, 123–132. doi: 10.1179/2050572813Y.0000000033

Rudner, M., Ng, E. H., Ronnberg, N., Mishra, S., Ronnberg, J., Lunner, T., et al. (2011). Cognitive spare capacity as a measure of listening effort. *J. Hear. Sci.* 1, 47–49.

Sacks, H., Schegloff, E. A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language (Baltim)* 50, 696. doi: 10.2307/412243

Sánchez-Cubillo, I., Periáñez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., et al. (2009). Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *J. Int. Neuropsychol. Soc.* 15, 438–450. doi: 10.1017/S1355617709090626

Shinn-Cunningham, B. G. (2005). Influences of spatial cues on grouping and understanding sound. *Proc. Forum Acusticum.*

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends Cogn. Sci.* 12, 182–186. doi: 10.1016/j.tics.2008.02.003

Singh, G., Pirhora-Fuller, M. K., and Schneider, B. A. (2008). The effect of age on auditory spatial attention in conditions of real and simulated spatial separation. *J. Acoust. Soc. Am.* 124, 1294–1305. doi: 10.1121/1.2949399

Singh, G., Pichora-Fuller, M. K., and Schneider, B. A. (2013). Time course and cost of misdirecting auditory spatial attention in younger and older adults. *Ear Hear.* 34, 711–721. doi: 10.1097/AUD.0b013e31829bf6ec

Sörqvist, P. (2010). The role of working memory capacity in auditory distraction: a review. *Noise Heal.* 12, 217–224. doi: 10.4103/1463-1741.70500

Sörqvist, P., Marsh, J. E., and Nöstl, A. (2013). High working memory capacity does not always attenuate distraction: bayesian evidence in support of the null hypothesis. *Psychon. Bull. Rev.* 20, 897–904. doi: 10.3758/s13423-013-0419-y

Sörqvist, P., and Rönnberg, J. (2014). Individual differences in distractibility: an update and a model. *PsyCh J.* 3, 42–57. doi: 10.1002/pchj.47

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10587–10592. doi: 10.1073/pnas.0903616106

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J. Speech Hear. Res.* 28, 455–462. doi: 10.1044/jshr.2803.455

Teder-Salejarvi, W. A., and Hillyard, S. A. (1998). The gradient of spatial auditory attention in free field: an event-related potential study. *Percept. Psychophys.* 60, 1228–1242. doi: 10.3758/BF03206172

Tun, P. A., McCoy, S., and Wingfield, A. (2009). Aging, hearing acuity, and the attentional costs of effortful listening. *Psychol. Aging* 24, 761–766. doi: 10.1037/a0014802

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *J. Acoust. Soc. Am.* 91, 1707–1717. doi: 10.1121/1.402450

Zurek, P. M. (1993). "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, eds G. A. Studebaker and I. Hochberg (Boston, MA: Allyn and Bacon), 255–276.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Lin and Carlile. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 07 May 2015
doi: 10.3389/fpsyg.2015.00526

Working memory and intelligibility of hearing-aid processed speech

Pamela E. Souza[1,2]*, *Kathryn H. Arehart*[3], *Jing Shen*[1], *Melinda Anderson*[3] and *James M. Kates*[3]

[1] Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA, [2] Knowles Hearing Center, Northwestern University, Evanston, IL, USA, [3] Department of Speech, Language and Hearing Sciences, University of Colorado at Boulder, Boulder, CO, USA

Previous work suggested that individuals with low *working memory capacity* may be at a disadvantage in adverse listening environments, including situations with background noise or substantial modification of the acoustic signal. This study explored the relationship between patient factors (including working memory capacity) and intelligibility and quality of modified speech for older individuals with sensorineural hearing loss. The modification was created using a combination of hearing aid processing [wide-dynamic range compression (WDRC) and frequency compression (FC)] applied to sentences in multitalker babble. The extent of signal modification was quantified via an envelope fidelity index. We also explored the contribution of components of working memory by including measures of processing speed and executive function. We hypothesized that listeners with low working memory capacity would perform more poorly than those with high working memory capacity across all situations, and would also be differentially affected by high amounts of signal modification. Results showed a significant effect of working memory capacity for speech intelligibility, and an interaction between working memory, amount of hearing loss and signal modification. Signal modification was the major predictor of quality ratings. These data add to the literature on hearing-aid processing and working memory by suggesting that the working memory-intelligibility effects may be related to aggregate signal fidelity, rather than to the specific signal manipulation. They also suggest that for individuals with low working memory capacity, sensorineural loss may be most appropriately addressed with WDRC and/or FC parameters that maintain the fidelity of the signal envelope.

Keywords: aging, cognition, hearing loss, hearing aid, compression, quality, intelligibility

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Staffan Hygge,
University of Gävle, Sweden
Rolph Houben,
Academic Medical Center
Amsterdam, Netherlands

***Correspondence:**
Pamela E. Souza,
Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60201, USA
p-souza@northwestern.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *01 February 2015*
Accepted: *13 April 2015*
Published: *07 May 2015*

Citation:
Souza PE, Arehart KH, Shen J, Anderson M and Kates JM (2015) Working memory and intelligibility of hearing-aid processed speech. Front. Psychol. 6:526. doi: 10.3389/fpsyg.2015.00526

Introduction

Individuals with hearing loss must frequently communicate under adverse conditions, including noisy, reverberant, or otherwise distorted speech. The ability to communicate in adverse listening environments is reduced by hearing loss, or when the individual is older (e.g., Pichora-Fuller and Souza, 2003). More recently, it has been proposed that individuals with low *working memory capacity* may also be at a disadvantage in adverse listening environments. Working memory capacity refers to the ability to simultaneously process and store information (Baddeley, 1992). During speech perception, listeners must extract meaning from acoustic patterns and store that meaning for integration with the ongoing auditory stream. When acoustic patterns are degraded or altered

from their expected form, it may be more difficult to match those acoustic patterns to stored lexical information (Rönnberg et al., 2013), and working memory may be engaged to a greater extent.

In the working memory model outlined by Baddeley (2000), the component of executive function (i.e., central executive) was included as the most important part of the working memory system. Its role was thought to be supervising, planning, and activating intentional actions. Other researchers' work illustrated this view more explicitly and defined executive function as shifting, updating, and inhibition in information processing (Miyake et al., 2000). In addition, speed of processing simple information was linked to working memory capacity in both older adults and children (Salthouse, 1991, 2000; Fry and Hale, 1996). These researchers proposed that individual difference in working memory capacity might be mediated by processing speed. Following from this idea, executive function and processing speed may also be related to signal modification in adverse listening conditions, consistent with the Ease of Language Understanding model (Rönnberg et al., 2013).

A common example of signal modification is speech in background noise. Everyday signal-to-noise ratios range from about +15 dB to as poor as −10 dB, with the most adverse situations including conversations in restaurants, automobiles, and public transportation (Olsen, 1998; Hodgson et al., 2007; Smeda et al., 2015). Listeners with low working memory capacity have more difficulty recognizing speech in noise than listeners with high working memory capacity (see Akeroyd, 2008 and Besser et al., 2013 for reviews). The association is stronger between verbal working memory tests and sentence intelligibility; and weaker between non-verbal working memory tests and syllables (e.g., Humes and Floyd, 2005). Moreover, some studies have shown a stronger relationship between working memory and sentence intelligibility when the sentences are presented at conversational or weaker levels to individuals with hearing loss (Humes and Floyd, 2005); or when the sentences are presented in modulated rather than unmodulated background noise (e.g., George et al., 2007). Presumably, both scenarios increase the number of inaudible or partially audible phonemes and the overall difficulty of the task, engaging working memory to a greater extent. The data on working memory capacity and speech in noise, then, are broadly consistent with the Rönnberg model.

While there are a large number of studies which measured working memory for speech in background noise, less information is available regarding other types of signal modification. For listeners with hearing loss, a potential source of signal modification is the signal processing applied by hearing aids. Only two decades ago, hearing aids were simple amplifiers where gain was dictated by the extent of hearing loss at each frequency, plus some means of limiting maximum output. Today, even "entry-level" hearing aids feature multiple features which may significantly modify the speech signal. Those features may include multichannel compression and output limiting, noise reduction, feedback suppression, and adaptive microphone directionality. Each feature has potential to alter the signal in a manner which may have consequences for the listener.

To illustrate this idea, consider wide-dynamic range compression (WDRC). WDRC is a core feature of digital hearing aids by which time-varying gain is applied to improve audibility of weak sounds while maintaining loudness comfort for higher-intensity sounds. The acoustic consequences of WDRC are dictated, in part, by the speed of the gain adjustment (i.e., attack and release times). In theory, fast compression which increases gain for brief speech segments will achieve greater consonant audibility than slow compression (e.g., Jenstad and Souza, 2005), and such compression is implemented in many commercial products. However, there is also evidence that alteration of the speech amplitude envelope—as will occur with fast compression (Kates, 2008)—may create a type of adverse listening situation for listeners who rely on envelope cues. A number of studies support the idea that listeners with low working memory capacity perform better with slow-acting than with fast-acting WDRC (e.g., Gatehouse et al., 2006; Lunner and Sundewall-Thoren, 2007; Davies-Venn and Souza, 2014; Ohlenforst et al., 2014; Souza and Sirow, 2014). Those data have been interpreted as a greater susceptibility to signal modification with low working memory capacity, which offsets the expected benefits of improved consonant audibility.

If susceptibility to signal modification is related to working memory capacity, we would expect to see similar patterns for other types of hearing-aid processing. One such example is frequency compression (FC). For listeners with substantial high-frequency loss, high-frequency gain may not result in audibility, either because gain is limited by the electroacoustic characteristics of the device, or because the listener may not have sufficient receptor cells to receive the amplified high-frequency cues (Moore, 2004). In FC, signal energy at high frequencies is digitally compressed into a lower frequency region where the listener has better hearing acuity. As with WDRC, the intent is to improve signal audibility. However, as with fast-acting WDRC, improved audibility requires signal modification. FC alters harmonic spacing and modifies spectral peak levels (McDermott, 2011). If the benefits of FC outweigh the (potential) disadvantage of such modification, speech intelligibility may be improved by signal modification (e.g., Souza et al., 2013; Alexander et al., 2014; Ellis and Munro, 2015). However, FC which results in extensive signal modification could also be viewed as creating an adverse listening environment for some listeners. Recent data show that the benefit of FC is influenced by working memory capacity, as well as age and amount of hearing loss (Arehart et al., 2013a; Kates et al., 2013). As with fast-acting WDRC, the FC data can be interpreted to show that listeners with low working memory capacity have greater susceptibility to signal modification caused by hearing-aid processing.

Although varying a single hearing-aid parameter is a reasonable way to model (potential) adverse listening situations for hearing-aid wearers, such implementations may not generalize to wearable hearing aids in which multiple parameters interact with (and perhaps offset) one another. We know that when signal processing algorithms are combined, speech intelligibility and quality ratings are different than when the algorithms process the same speech in isolation (e.g., Franck et al., 1999; Chung, 2007; Anderson et al., 2009). Related to working memory, recent work by Neher and colleagues (Neher et al., 2013, 2014; Neher, 2014) explored the relationship between working memory, executive function, and response to aggregate signal modification. In

Neher's work, signal modification was created by a combination of background noise, hearing aid noise reduction and directional microphones. The extent of signal modification was manipulated by controlling the level of background noise and/or the strength of the noise reduction algorithm. Consistent with (Arehart et al., 2013b), more aggressive noise reduction was verified to result in greater signal modification. In agreement with previous work for other types of hearing aid processing, working memory capacity and amount of hearing loss predicted amplified speech intelligibility.

To summarize, a growing body of work suggests that a relationship between working memory capacity and listening in adverse conditions can be demonstrated not only for environmental distortions such as background noise (Akeroyd, 2008), but for signal modification introduced by hearing devices. In this study, we explored the relationship between signal modification, speech intelligibility, and working memory capacity, where signal modification was the aggregate effect of background noise and simulated amplification with two processing strategies: amplitude compression, and FC. Each strategy was further manipulated by applying parameters which would modify the signal to a greater or lesser extent. Here, we hypothesize that signal modification created by amplification is related to working memory capacity, such that the resulting modification is the key factor. If that holds true, it would be consistent with Rönnberg and colleagues' model of working memory (Rönnberg et al., 2013), in which greater modification of the expected acoustic signal places a greater demand on working memory capacity. Participants were older adults with mild-to-moderate hearing loss. Working memory capacity was quantified using a reading span test (RST). Executive function and processing speed were also measured in order to evaluate their relationship to intelligibility of speech. We posed three questions: (1) How does the performance of speech intelligibility (and quality) vary across adverse listening conditions? (2) What role do listener factors such as cognitive ability, amount of hearing loss, and age have in speech intelligibility (and quality) performance under such adverse listening conditions? (3) Is there a cognitive factor (specifically, working memory capacity, executive function, or processing speed) that improves prediction of intelligibility in adverse listening conditions?

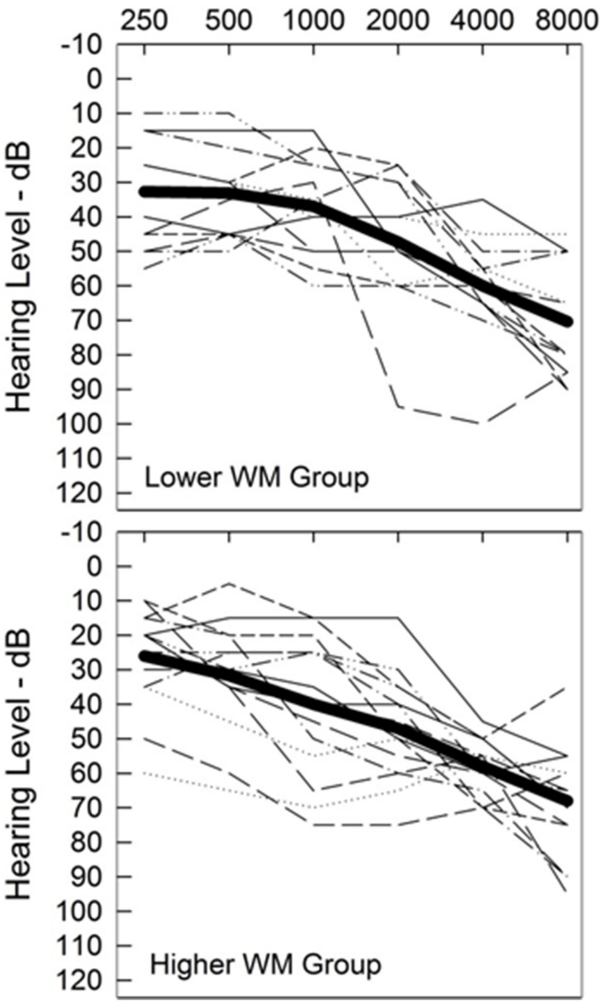


FIGURE 1 | Individual test-ear audiograms (thin lines). Audiograms are grouped by working memory capacity (described in detail later in this paper). Audiograms for participants with lower working memory capacity (WM) are shown in the top panel and with higher working memory in the bottom panel. The average audiogram for each group is shown with a heavy line.

Materials and Methods

Participants

Participants were recruited and data collected across two study sites (Northwestern University and University of Colorado), using identical test equipment and protocols. Twenty-nine older participants aged 49–89 years (mean age 74.0 years) participated in the study. Inclusion criteria included symmetrical sensorineural hearing loss with thresholds between 25 and 70 dB HL at octave frequencies between 0.5 and 3 kHz; a difference in pure-tone average [0.5, 1, 2 kHz] $\leq$ 10 dB across ears; and air-bone gaps $\leq$10 dB. One ear was randomly selected as the test ear for the auditory portions of the study. Test ear thresholds are shown in **Figure 1**, grouped by working memory capacity (explained in detail later in this paper). Quiet word recognition scores (monosyllabic words presented to the test ear at 30–40 dB SL) ranged from 68 to 100% (mean score 88%). All participants had good self-reported health, normal or corrected-to-normal vision, and completed a cognitive screening using the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005). This brief (10 min) cognitive screening test assesses attention, working memory, executive function, visual-spatial ability, and language skills. Participants scoring 22 or higher on the MoCA were accepted into the study. That inclusion criterion considered the effects of hearing loss (Dupuis et al., 2013) and participant demographics (Rosetti et al., 2011), and was similar to that followed in previous studies with the same population (Anderson et al., 2012, 2013). Testing (audiometric evaluation, speech intelligibility, quality ratings, working memory capacity, executive function, and processing speed) was completed over test sessions of 1–2 h

each, including test breaks. Ethical and safety review of the test protocol was conducted and approved by the local institutional review board at each site. Participants were compensated for their time.

Working Memory Test

The RST (Daneman and Carpenter, 1980; Rönnberg et al., 1989) was used to measure working memory. The test was designed to measure individual working memory capacity in terms of coordinating storage and processing requirements simultaneously. During the test, 54 sentences were shown on the computer screen one word or word pair at a time (on-screen duration 800 ms). Half of the sentences were absurd (e.g., "The train" "sang" "a song"), and half were semantically meaningful (e.g. "The captain" "sailed" "his boat"). The participants were asked to read each sentence and make a semantic judgment as to the sense of the sentence. After each 3–6 sentence block, the participants were asked to recall the first or the last words of a presented set of sentences. The primary measure of the individual's working memory capacity was the proportion of words that were correctly recalled.

Processing Speed and Executive Function

The flanker task (Eriksen and Ericksen, 1974) was used to measure the participants' processing speed and executive function. In this task, the participants were asked to identify the direction of an arrow that was presented on the center of the screen. Processing speed was quantified by reaction time (in milliseconds) to a single arrow on the screen without any visual interference. Executive function was quantified by the difference in reaction time when the central arrow was flanked by arrows that had the same (congruent) vs. different (incongruent) directions as the center arrow.

The participants were seated in front of a computer monitor with eye-to-screen distance of 17 inches. They were asked to press the button corresponding to the direction of the arrow (i.e., press left button when the arrow pointed left, press right button when the arrow pointed right) as quickly and as accurately as possible. A practice block (8 trials for the processing speed test, 12 trials for the executive function test) was conducted prior to each test in order to ensure the instruction was followed. The processing speed test had one block of 40 trials. The arrow was pointing left in half of the trials and pointing right in the other half. The executive function test had one block of 80 trials. Three arrows on each side surrounded the center arrow in each trial. The side arrows were pointing to the same direction as the center arrow in half of the trials, while pointing a different direction in the other half. The order of the trials was randomized across participants.

Speech Intelligibility and Quality Stimuli

Speech intelligibility and quality were measured using materials drawn from the Institute of Electrical and Electronics Engineers sentence corpus (Rosenthal, 1969). This corpus consists of a large set of sentences which make semantic sense but contain relatively little contextual information. Each sentence includes five key words which can be scored for correct repetition (e.g., "The birch canoe slid on the smooth planks"; "Glue the sheet to the dark blue background."). The sentences were spoken by a female talker and were digitized at a 44.1 kHz sampling rate and then downsampled to 22.05 kHz. The level of the sentences at the input to the hearing-aid simulation was set at 65 dB SPL. The final presentation level was based on the individualized frequency-gain shaping described below.

To create realistic adverse listening conditions, the sentences were digitally combined with multi-talker babble (Cox et al., 1987) at two signal-to-noise ratios, 0 and +10 dB, plus a quiet (no noise) condition. For each signal-to-noise ratio, the sentences were set to a level of 65 dB SPL and the noise level adjusted prior to mixing.

Hearing Aid Processing

Dynamic-range compression (WDRC) was implemented using a hearing aid simulation program with 6-channel FIR filter bank. The center frequencies of the bands were 250, 500, 1000, 2000, 4000, and 6000 Hz. Inputs having intensities below a lower compression threshold (45 dB SPL) received linear amplification, and inputs above an upper compression threshold (100 dB SPL) received compression limiting to prevent over-amplification of intense sounds. Input levels between the two compression thresholds were subjected to WDRC with a compression ratio of 2:1. There were two WDRC conditions, with release times of 40 and 640 ms (re: ANSI, 2009). The attack time was set to 5 ms in both cases. In a control condition, linear processing was implemented using the same algorithm, but with the compression ratio set to 1:1.

FC was implemented using sinusoidal modeling (McAulay and Quatieri, 1986). The signal was separated into two frequency bands above and below each of the cutoff frequencies specified below. The low-frequency band was used without processing, while FC was applied to the high-frequency band using short-time frequency analysis, as follows: (1) the high-frequency signal was windowed in 6 ms segments using a von Hann raised-cosine window; (2) the shifted frequency components used the original amplitude and phase values, applied to sinusoids generated at the new frequencies; (3) the synthesized high-frequency and original low-frequency signals were recombined in the final step to produce the processed output.

Two FC conditions were used to present strong and mild signal modification (Strong: FC cutoff of 1000 Hz, FC ratio of 3:1; Mild: FC cutoff of 1500 Hz, FC ratio of 1.5:1). There was also a control condition with no FC applied to the signal.

To accommodate the individual hearing losses, all processed stimuli were amplified using the National Acoustics Laboratories-Revised (NAL-R) linear prescriptive formula (Byrne et al., 2001) with the gain implemented using a 128-point linear-phase FIR digital filter.

Signal Fidelity

Signal modifications to the original speech signal caused by cumulative effects of the additive noise and the signal processing were quantified using a signal fidelity metric (Kates and Arehart, 2014). The metric starts with an auditory model that reproduces the fundamental aspects of the auditory periphery including auditory frequency analysis, the dynamic-range compression

mediated by the outer hair cells, firing-rate adaptation associated with the inner hair cells, and auditory threshold. The output of the auditory model is the speech envelope in 32 auditory frequency bands from 80 to 8000 Hz.

The envelope outputs from the model for an unmodified reference signal having no noise or distortion are compared to the model outputs for the degraded signal. At each time sample, a smoothed version of the auditory spectrum is formed. The variations as a function of time in the smoothed spectrum for the modified signal are compared to the variations in the reference signal using a normalized cross-correlation operation. The resultant metric thus combines (1) the accuracy in reproducing the short-time spectral shape across auditory bands and (2) the accuracy in reproducing the envelope temporal modulation within auditory bands. The metric therefore provides an overall measure of fidelity in reproducing the time-frequency modulation pattern of the modified signal in a manner consistent with the time-frequency modulation patterns of speech (Zahorian and Rothenberg, 1981). The metric values range from 0 to 1, with 0 indicating a complete lack of envelope fidelity relative to the reference and 1 indicating perfect envelope fidelity.

Speech Intelligibility

For the intelligibility tests, the participant was seated in a double-walled sound booth and listened to stimuli presented monaurally via a Sennheiser HD 25 1 II headphone in the better ear. Each trial consisted of a sentence randomly drawn from one of the 27 processing conditions (3 WDRC × 3 FC × 3 signal-to-noise ratios). Subjects first heard 27 practice sentences (1 from each test condition) and then listened to 270 test sentences (with 10 sentences in each condition). No feedback was provided. The timing of presentation was controlled by the participant. The participant repeated the sentence and scoring was completed by the experimenter, seated outside the sound booth. The order of sentences and conditions was randomized across listeners. Scores were calculated based on the proportion of correctly-identified key words (10 sentences per condition and 5 words per sentence for 50 key words per condition, per participant).

Speech Quality

In the speech quality task, listeners rated the sound quality of speech that had been modified according to processing conditions discussed above. Stimuli were spoken by a woman, and were two sentences taken form the IEEE corpus ("Take the winding path to reach the lake. A saw is a tool used for making boards"). Each trial included the same two sentences to limit the effects of intelligibility. Speech processed by hearing aid signal processing algorithms have been shown to be well predicted by metrics using a single "overall quality" rating scale (e.g., Arehart et al., 2010), even though sound quality is multidimensional in nature (Gabrielsson et al., 1988; Arehart et al., 2007). In this study, listeners used a computer-based slider bar to rate the sound quality using a rating scale from 0 (poor sound quality) to 10 (excellent sound quality) in 0.5 increment (ITU, 2003[1]). The participant controlled the timing of presentation. Testing was completed in four blocks. The first block was a practice block, and included one trial from each of the processing conditions. The practice block familiarized the listener with the task and process of using the rating scale. Three test blocks followed, with 45 trials per block. Processing conditions were presented five times each, and were randomized to occur at any point within the three test blocks. No feedback was provided.

[1]International Telecommunication Union ITU-R: BS.1284-1, "General Methods for the Subjective Assessment of Sound Quality" (2003).

Results

Working Memory

Individual working memory scores are plotted in **Figure 2** as a function of amount of hearing loss (pure-tone average for 0.5, 1, 2 kHz). Scores ranged from 15 to 54%, with a mean score of 38%. The distribution of scores was similar to scores in previous studies which used the same reading span implementation, and where mean reading span scores ranged from 34 to 44% (e.g., Foo et al., 2007; Arehart et al., 2013a,b; Souza and Sirow, 2014). Within our test cohort there was no relationship between working memory capacity and amount of hearing loss ($r = -0.045$, $p = 0.817$). For some of the planned analyses, the participants were assigned to either a high ($n = 13$) or low ($n = 16$) working memory group, based on the median score for the group. Individuals who fell on the median were assigned to the higher group. Those groupings are indicated by different symbols in **Figure 2**.

Statistical Analysis

Similar to other work from our group (e.g., Arehart et al., 2013a), the primary analytical approach was hierarchical linear modeling (HLM) also known as multi-level modeling (Singer and Willett, 2003). Multi-level models were developed for the analysis of nested data structures or repeated measures data. They

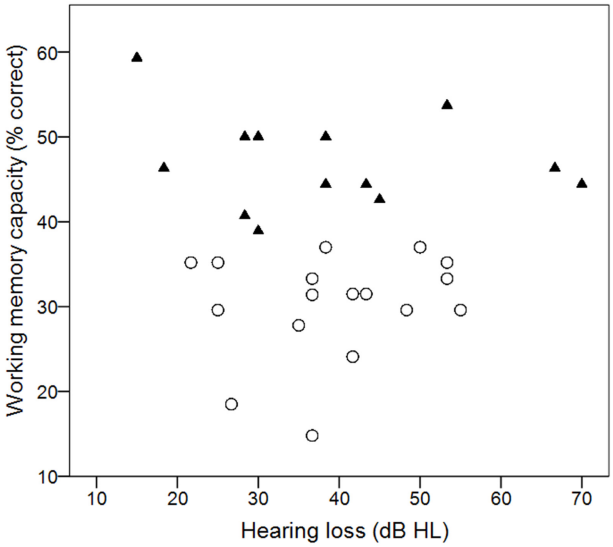


FIGURE 2 | Individual working memory scores as a function of hearing loss. Filled triangles and open circles show individual scores that fall above or below the median score.

incorporate between-listener characteristics in models of individual performance across multiple conditions (Raudenbush and Bryk, 2002), so are well suited for research questions where the variability in outcomes may be a result of differences between groups as well as individual listener differences.

The analysis was conducted using HLM 6 (Raudenbush and Bryk, 2002) and included three different multi-level models. Each model considered signal modification (using the envelope fidelity metric described above), amount of hearing loss (expressed as the average of thresholds at 1, 2, 3, and 4 kHz in the test ear) and age; plus one of the cognitive measures (working memory capacity, executive function, or processing speed). Listeners were grouped for amount of hearing loss, working memory capacity, executive function, and processing speed using the median as the cutoff criteria. Individuals who fell on the median were assigned to the higher scoring group.

Speech Intelligibility

Figures 3, **4** show mean intelligibility scores for each processing condition, grouped by working memory capacity. Recall that signal modification was created by manipulating three aspects of the signal: the amount of background noise; the WDRC release time; and the FC parameters. In **Figure 3**, data are plotted for the three WDRC conditions (collapsed across FC). In **Figure 4**, data are plotted for the three FC conditions (collapsed across WDRC). Each panel shows a different signal-to-noise ratio. Several trends are apparent. Scores were lower with more background noise; with more aggressive FC; and with faster WDRC (although the latter difference was quite small and occurred only at the poorest signal-to-noise ratio). With regard to working memory capacity, listeners with higher working memory performed better than their counterparts with low working memory across all conditions.

The rationale for the various background noise levels and the WDRC and FC processing was to create a range of signal modification, which was expected to underlie intelligibility (and perhaps quality) results. **Figure 5** shows average intelligibility scores as a function of the envelope fidelity metric. The envelope fidelity metric was subjected to a sigmoidal transformation to better support the model's assumption of linearity prior to HLM analysis. Each processing combination is indicated by data point labeling, and signal-to-noise ratio is indicated by symbol shape/color. Overall, there was a strong linear relationship between speech intelligibility and the (transformed) fidelity metric ($R^2 = 0.88$).

Model Fit and Definitions

The multi-level model for this analysis had two levels. The first level represented the individual linear relationship between speech intelligibility and envelope fidelity using estimated intercepts and slope coefficients. Listeners were then classified into groups based on their individual characteristics as described in the analysis section. Those groupings represented the model's second level, where listener characteristics were used to predict variability in the level one coefficients of intercept and slope. If un-centered, the intercept coefficient would have represented speech intelligibility at an envelope fidelity value of zero, where

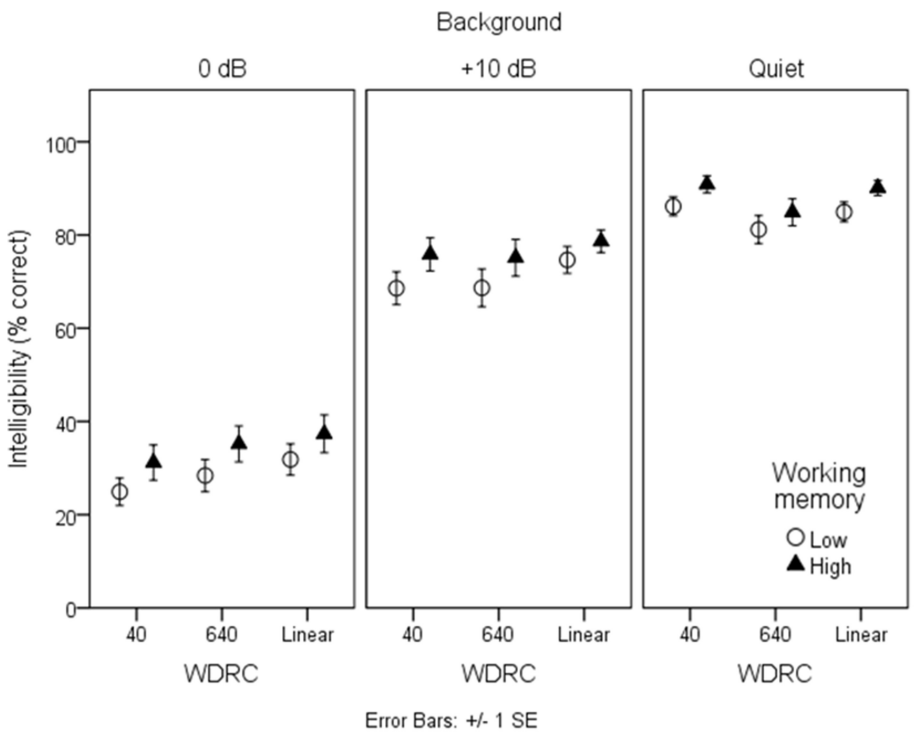


FIGURE 3 | Mean intelligibility for low- and high-working memory groups by WDRC condition. Error bars show ± one standard error about the mean.

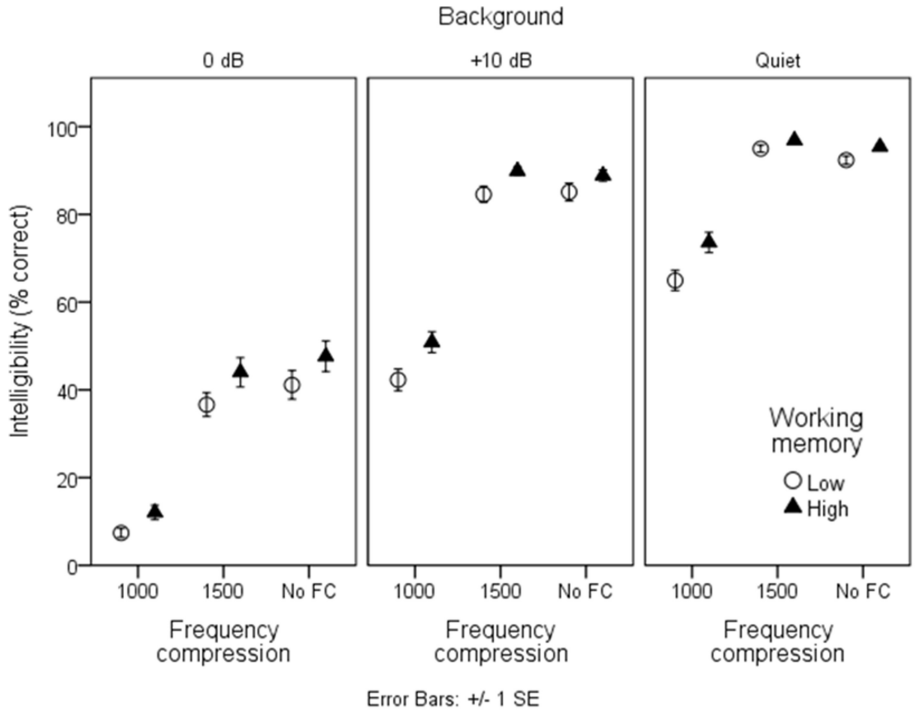


FIGURE 4 | As in Figure 3, but grouped by frequency compression condition.

signal modification was very high with minimal between-group differences. Accordingly, we centered the intercept at the mean of the envelope fidelity scale. Centering at the mean of the scale provided a more informative estimation of between group differences.

Between-listener Variability and Descriptive Statistics

The average estimated intelligibility for intercept across all listeners and conditions was 63.5% ($SD = 9\%$) and the average estimate for slope was 1 ($SD = 0.08$). To get a reference as to the magnitude of between-group differences in intercept and slope, we calculated the predicted 95% range for each coefficient. The predicted range for speech intelligibility intercept was 45.84 to 81.14% and the range for slope was 0.84 to 1.16. Recall that to predict between-listener variability, we explored a hierarchy of conditional models for each cognitive measure (working memory, executive function and processing speed).

Working memory scores (in proportion correct) ranged from 0.19 to 0.59, with a mean score of 0.38. The average processing speed score was 478 ms (range 361 to 606 ms). The average executive function score was 46 ms (range −64 to 204 ms). Correlations between the three cognitive measures (**Table 1**) were low and were not significant, suggesting that the three measures represented different cognitive domains.

Hierarchical Linear Model

The HLM model building process included predictors stepwise in an effort to partial out the amount of variability explained as well as the effect size for different listener factors. In each model, the first step included one of the three cognitive measures. In step 2 amount of hearing loss was added, followed by age in the third step.

Table 2 provides a summary of the fixed effects for the working memory model hierarchy. In step 1 the results show that there was a significant positive effect for envelope fidelity on speech intelligibility ($p < 0.001$). However there was no main effect for working memory capacity on intercept or slope. In step 2, when amount of hearing loss (pure-tone average, PTA) was added to the model, there was a significant main effect for working memory capacity ($p = 0.032$) and amount of hearing loss ($p < 0.001$) on intercept but no effect for either factor on slope. In other words, after controlling for amount of hearing loss there was a significant difference in speech intelligibility between the high and low working memory groups when envelope fidelity was at the mean of its scale. In step 3, age was added to the model but did not demonstrate any significant effects.

The change in the effect of working memory with the addition of amount of hearing loss indicated the presence of an underlying interaction. In step four, we removed age from the model and added a three way interaction (working memory by amount of hearing loss by envelope fidelity). The results of the final model demonstrated significant effects for working memory capacity ($p = 0.032$) and amount of hearing loss ($p < 0.001$) on intercept and a significant effect for working memory ($p = 0.005$) on slope. There was also a significant main effect for the three way interaction on speech intelligibility ($p = 0.011$).

Tables 3, **4** provide the model outcomes when executive function and processing speed were considered the primary cognitive

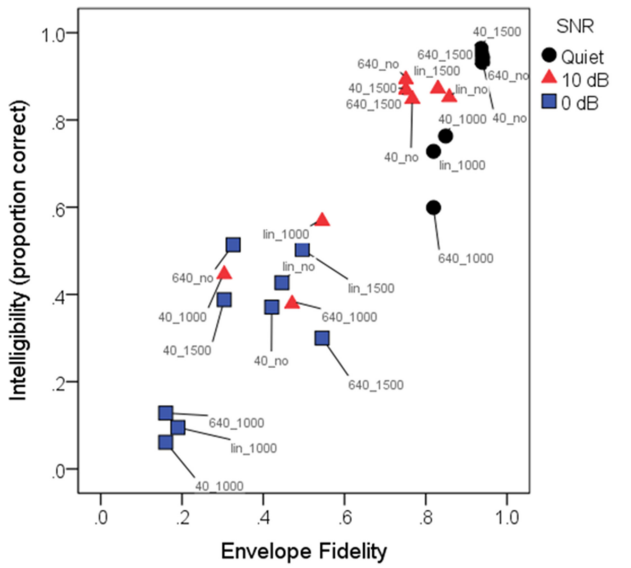


FIGURE 5 | Mean intelligibility scores (in proportion correct) as a function of envelope fidelity. For linearity prior to analysis, the envelope distortion metric was subjected to a sigmoidal transformation. Pearson product-moment correlation was 0.93, indicating that the envelope fidelity metric was a good predictor of intelligibility scores. Each condition is indicated by a color-label combination. Symbols indicate the three SNRs: quiet (black circles); 10 dB SNR (red triangles); 0 dB (blue squares). Labels indicate the conditions, where 40_, 640_, and lin_ represent WDRC release time or linear amplification, and _1000, _1500, or _no represent frequency compression cutoff frequency or no FC. As quantified by the envelope fidelity metric, the highest-fidelity condition was linear amplification without frequency compression for speech in quiet, and the lowest-fidelity condition was speech at a signal-to-noise ratio of 0 dB with a compression release time of 40 ms and a frequency compression cutoff of 1000 Hz. Note that due to the close clustering of symbols with high (near-100%) intelligibility, not all symbols and labels are visible in the figure.

TABLE 1 | Pearson product-moment correlations between cognitive measures.

	Processing speed (ms)	Executive Function (ms)	Working memory (% correct)
Processing speed (ms)	1.00	−0.07	0.16
Executive function (ms)		1.00	−0.10
Working memory capacity (% correct)			1.00

TABLE 2 | Summary of hierarchical linear model for intelligibility with working memory capacity (WM) as a predictor.

Fixed effect	Coeff.	Std. error	T-ratio	*d.f*	*P*-value
STEP 1					
For Intercept					
Intercept	0.604	0.026	23.675	27	<0.001
WM	0.061	0.033	1.860	27	0.073
For Slope					
Fidelity index	1.048	0.023	45.665	27	<0.001
WM	−0.070	0.042	−0.649	27	0.110
STEP 2					
For Intercept					
Intercept	0.603	0.023	26.132	26	<0.001
WM	0.063	0.028	2.269	26	0.032
PTA	−0.003	0.001	−5.935	26	<0.001
For Slope					
Fidelity index	1.049	0.024	42.915	26	<0.001
WM	−0.071	0.041	−1.733	26	0.094
PTA	0.002	0.001	1.815	26	0.081
STEP 3					
For Intercept					
Intercept	0.606	0.021	28.792	25	<0.001
WM	0.056	0.027	2.115	25	0.044
PTA	−0.003	0.0004	−7.001	25	<0.001
Age	−0.001	0.001	−0.881	25	0.387
For Slope					
Fidelity index	1.052	0.026	40.404	25	<0.001
WM	−0.076	0.042	−1.774	25	0.088
PTA	0.002	0.001	−1.820	25	0.080
Age	0.000	0.002	−0.325	25	0.748
STEP 4					
For Intercept					
Intercept	0.737	0.036	20.26	26	<0.001
WM	0.063	0.027	2.269	26	0.032
PTA	−0.003	0.001	−5.935	26	<0.001
For Slope					
Fidelity index	1.090	0.066	16.560	25	<0.001
WM	−0.265	0.086	−3.089	25	0.005
PTA	−0.001	0.001	−0.765	25	0.451
WM by PTA	0.005	0.002	2.747	25	0.011

Amount of hearing loss (PTA) is average of thresholds at 1, 2, 3, and 4 kHz in the test ear.

predictor. Neither of these factors was significant predictors of speech intelligibility, either independently or when controlling for amount of hearing loss and age.

Effect Sizes and Prototypical Plots

The working memory model represented in step 4 of **Table 2** explained 33% of variability in intercept and 21% of variability in slope. When controlling for amount of hearing loss, listeners in the higher working memory group had an estimated gain of 6.3% in intelligibility at the mean envelope fidelity. As expected, speech intelligibility scores decreased as envelope fidelity decreased. However, after controlling for amount of loss and the hearing loss-by-working memory interaction, listeners' scores in the high working memory group decreased at a slower rate (8.2% per fidelity unit) when compared to listeners in the low working memory group (10% per fidelity unit). Finally, the interaction demonstrated that as envelope fidelity decreased, listeners with milder hearing loss and high working memory capacity tended to have higher intelligibility scores compared to listeners with milder hearing loss and low working memory capacity. As hearing loss increased, the relationship between working memory and speech intelligibility diminished.

TABLE 3 | Summary of hierarchical linear model for intelligibility with executive function (EF) as a predictor.

Fixed effect	Coeff.	Std. error	T-ratio	*d.f*	*P*-value
STEP 1					
For Intercept					
Intercept	0.644	0.029	22.403	26	<0.001
EF	−0.026	0.035	−0.758	26	0.455
For Slope					
Fidelity index	1.025	0.033	31.014	26	<0.001
EF	−0.013	0.045	−0.286	26	0.777
STEP 2					
For Intercept					
Intercept	0.769	0.043	17.923	25	<0.001
EF	−0.018	0.030	−0.615	25	0.544
PTA	−0.003	0.001	−4.310	25	<0.001
For Slope					
Fidelity index	0.943	0.070	13.495	25	<0.001
EF	−0.018	0.042	−0.423	25	0.676
PTA	0.002	0.001	1.638	25	0.114
STEP 3					
For Intercept					
Intercept	0.951	0.098	9.708	24	<0.001
EF	−0.014	0.030	−0.461	24	0.649
PTA	−0.003	0.001	−6.048	24	<0.001
Age	0.003	0.002	−1.621	24	0.118
For Slope					
Fidelity index	0.879	0.193	4.556	24	<0.001
EF	−0.020	0.045	−0.439	24	0.664
PTA	0.002	0.001	1.687	24	0.104
Age	0.001	0.003	0.336	24	0.740

Amount of hearing loss (PTA) is average of thresholds at 1, 2, 3, and 4 kHz in the test ear.

TABLE 4 | Summary of hierarchical linear model for intelligibility with processing speed (PS) as a predictor.

Fixed effect	Coeff.	Std. error	T-ratio	*d.f*	*P*-value
STEP 1					
For Intercept					
Intercept	0.626	0.020	30.675	26	<0.001
PS	0.008	0.034	0.246	26	0.808
For Slope					
Fidelity index	1.029	0.033	30.624	26	<0.001
PS	−0.022	0.044	−0.503	26	0.619
STEP 2					
For Intercept					
Intercept	0.763	0.038	20.001	25	<0.001
PS	−0.003	0.029	−0.110	25	0.914
PTA	−0.003	0.001	−3.948	25	0.001
For Slope					
Fidelity index	0.946	0.072	13.105	25	<0.001
PS	−0.015	0.043	−0.353	25	0.726
PTA	0.002	0.001	1.523	25	0.140
STEP 3					
For Intercept					
Intercept	0.951	0.10	8.671	24	<0.001
PS	−0.001	0.027	−0.040	24	0.969
PTA	−0.003	0.001	−5.637	24	<0.00 1
Age	−0.003	0.001	−1.755	24	0.092
For Slope					
Fidelity index	0.886	0.190	4.665	24	<0.001
PS	−0.016	0.043	−0.371	24	0.713
PTA	0.002	0.001	1.548	24	0.134
Age	0.001	0.002	0.338	24	0.738

Amount of hearing loss (PTA) is average of thresholds at 1, 2, 3, and 4 kHz in the test ear.

To illustrate the simultaneous effects of all the predictors in the final model for RST, we created a model plot with prototypical listener characteristics. **Figure 6** illustrates the model for intelligibility in step 4 and provides four different fitted trajectories of intelligibility as a function of envelope fidelity. The fitted trajectories represented two subsets of listeners within the High and Low working memory groups. In the first subset hearing loss was modeled at the 25th percentile (28 dB HL pure-tone average) and for the second subset hearing loss was modeled at the 75th percentile (49 dB HL pure-tone average).

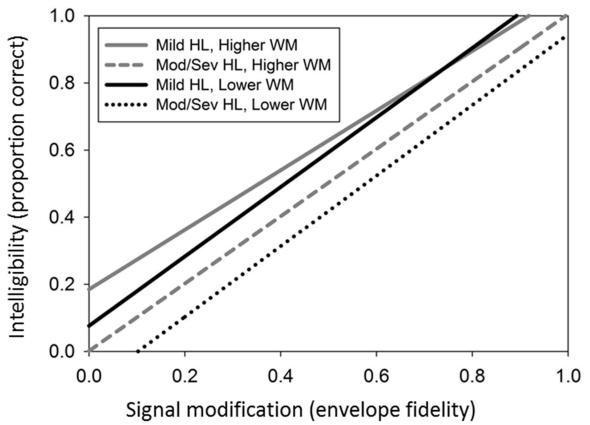


FIGURE 6 | Final model for intelligibility, showing four different fitted trajectories of intelligibility as a function of envelope fidelity for hearing loss (HL) and working memory (WM).

Speech Quality

Figures 7, **8** show mean quality ratings for each processing condition. For consistency with the intelligibility figures, listeners are grouped by working memory. In **Figure 7**, data are plotted for the three WDRC conditions (collapsed across FC). In **Figure 8**, data are plotted for the three FC conditions (collapsed across WDRC). Each panel shows a different signal-to-noise ratio. In contrast to the intelligibility data (**Figures 3**, **4**), there was no suggestion that working memory capacity influenced quality ratings in a consistent way. However, we anticipated that quality ratings would depend to a large extent on signal modification. **Figure 9** shows average quality ratings as a function of the envelope fidelity metric. Each processing combination is indicated by data point labeling, and signal-to-noise ratio is indicated by symbol shape/color.

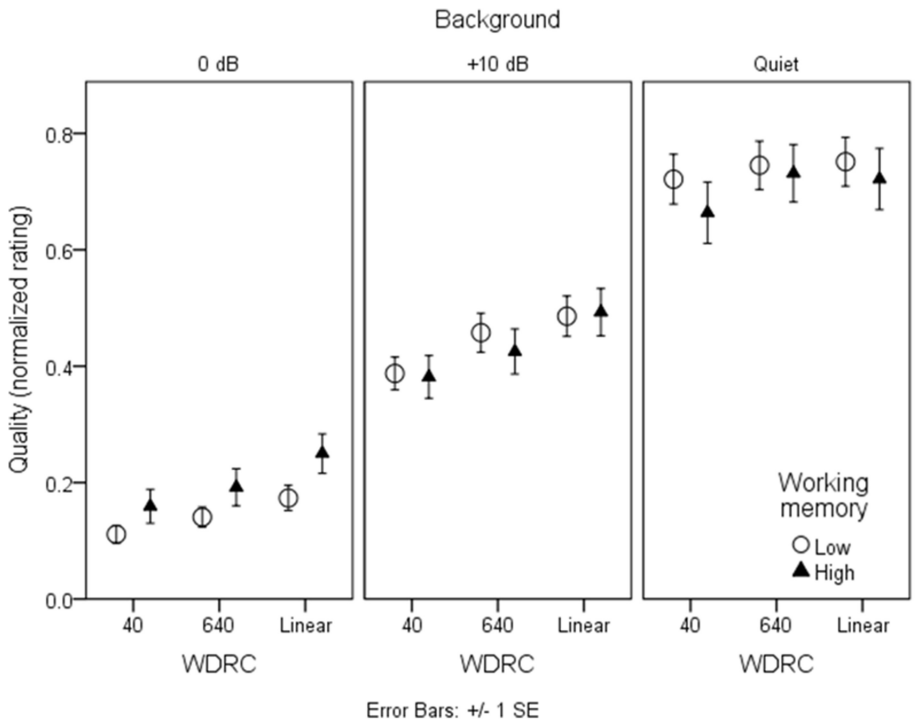


FIGURE 7 | As in Figure 3, but for quality ratings grouped by WDRC condition.

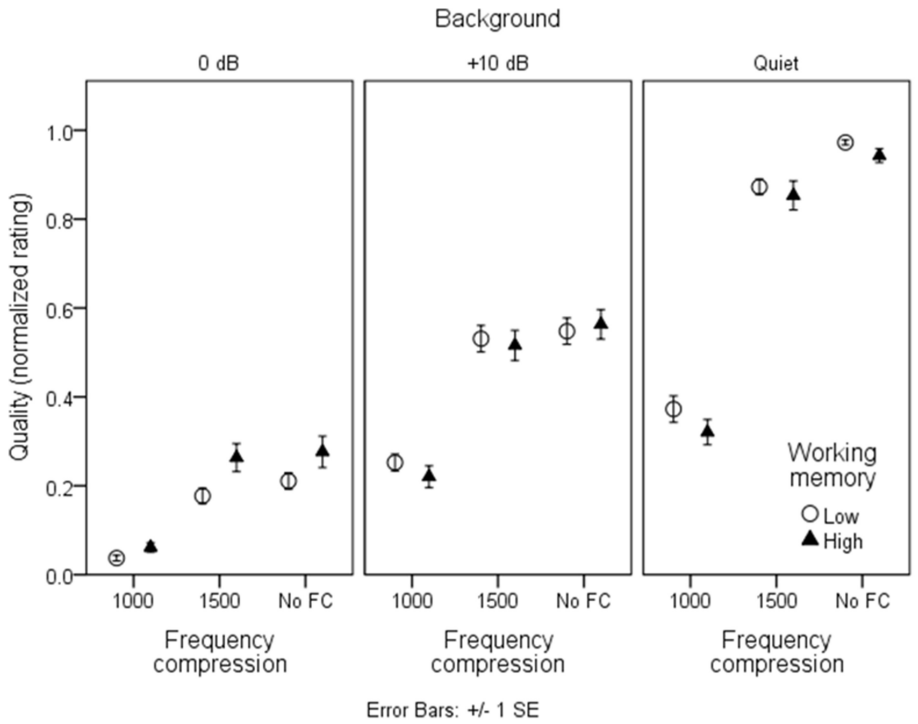


FIGURE 8 | As in Figure 3, but for quality ratings grouped by frequency compression condition.

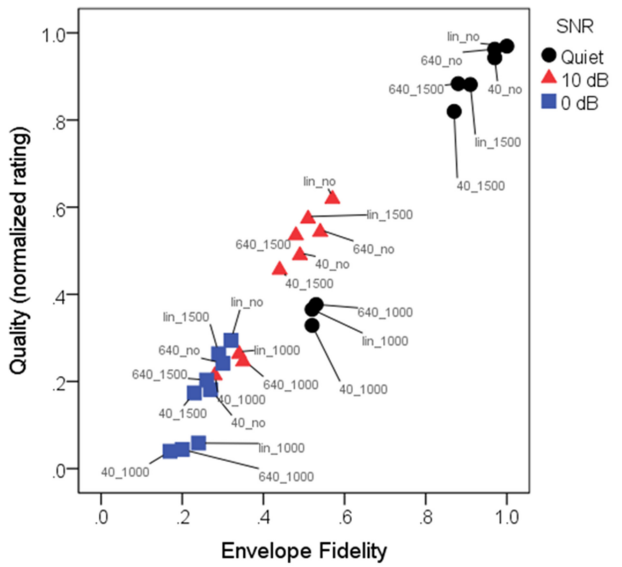


FIGURE 9 | Mean normalized quality ratings as a function of envelope fidelity. Pearson product-moment correlation was 0.97, indicating that the envelope fidelity measure was a good predictor of quality ratings. Each condition is indicated by a color-label combination. Symbols indicate the three SNRs: quiet (black circles); 10 dB SNR (red triangles); 0 dB (blue squares). Labels indicate the conditions, where 40_, 640_, and lin_ represent WDRC release time or linear amplification, and _1000, _1500, or _no represent frequency compression cutoff frequency or no FC. As quantified by the envelope fidelity metric, the highest-fidelity condition was linear amplification without frequency compression for speech in quiet, and the lowest-fidelity condition was speech at a signal-to-noise ratio of 0 dB with a compression release time of 40 ms and a frequency compression cutoff of 1000 Hz.

There was a strong linear relationship between speech quality and the fidelity metric ($R^2 = 0.88$).

Between-group Variability

The average estimate for quality intercept across all listeners and conditions was 0.44 ($SD = 0.08$) and the average estimate for slope was 1.1 ($SD = 0.14$). The predicted 95% range for quality intercept was 0.28 to 0.60 and the range for slope was 0.83 to 1.37.

Hierarchical Linear Model

Similar to the speech intelligibility analysis, we also included three HLM models for quality in order to identify the independent effect for each cognitive measure. The model building process included predictors stepwise where the first step included one of the three cognitive measures independently. The next step added PTA as a covariate and the third step added age also as a covariate to the model.

Tables 5–**7** summarize the parameter coefficients for each HLM model and sub-models provide for quality. The first level model demonstrated that there was a statistically significant effect for envelope fidelity ($p < 0.001$) on quality ratings. For the working memory model, we found no significant effects for working memory group, amount of hearing loss, or age. Similarly, there were no significant effects for processing speed group, amount of hearing loss, or age in the processing speed model (**Table 7**). The executive function model did reveal a small significant effect for executive function group and age on intercept.

TABLE 5 | Summary of hierarchical linear model for quality with working memory capacity (WM) as a predictor.

Fixed effect	Coeff.	Std. error	T-ratio	*d.f*	*P*-value
STEP 1					
For Intercept					
Intercept	0.446	0.019	22.759	27	<0.001
WM	−0.005	0.032	−0.144	27	0.887
For Slope					
Fidelity index	1.127	0.032	29.444	27	<0.001
WM	−0.056	0.065	−0.868	27	0.393
STEP 2					
For Intercept					
Intercept	0.461	0.043	10.817	26	<0.001
WM	−0.004	0.031	−0.129	26	0.899
PTA	−0.0007	0.001	−0.642	26	0.526
For Slope					
Fidelity index	1.014	0.107	9.480	26	<0.001
WM	−0.058	0.063	−0.918	26	0.367
PTA	0.003	0.003	1.154	26	0.259
STEP 3					
For Intercept					
Intercept	0.249	0.111	2.246	25	0.034
WM	0.011	0.031	0.360	25	0.721
PTA	−0.001	0.001	−0.572	25	0.572
Age	0.003	0.002	1.886	25	0.071
For Slope					
Fidelity index	1.164	0.295	3.944	25	0.001
WM	−0.069	0.063	−1.082	25	0.290
PTA	0.003	0.003	1.135	25	0.268
Age	−0.002	0.003	−0.572	25	0.572

Amount of hearing loss (PTA) is average of thresholds at 1, 2, 3, and 4 kHz in the test ear.

Discussion

Our first question concerned speech intelligibility (and quality) across adverse listening conditions. We considered "adverse" quite broadly to mean addition of background noise and/or modifications of the acoustic signal (here, by WDRC and FC). An envelope fidelity metric was used to quantify those modifications. Speech intelligibility and quality were well predicted by the envelope fidelity metric.

Next, we explored the role of listener factors on speech intelligibility (and quality) performance under adverse listening conditions. The patient factors that were considered were amount of hearing loss, age, working memory capacity, executive function and processing speed. The focus of the study was working memory capacity, which had already been shown to be related to hearing aid processing parameters when a single type of processing was applied. A recent model of working memory (Rönnberg et al., 2013) suggests that when signal modification impedes a rapid match of acoustic information to stored representations, working memory will be engaged. In that situation, listeners with low

TABLE 6 | Summary of hierarchical linear model for quality with executive function (EF) as a predictor.

Fixed effect	Coeff.	Std. error	T-ratio	*d.f*	*P*-value
STEP 1					
For Intercept					
Intercept	0.473	0.022	21.186	26	<0.001
EF	−0.058	0.031	−1.867	26	0.073
For Slope					
Fidelity index	1.081	0.039	27.858	26	<0.001
EF	0.028	0.067	0.415	26	0.681
STEP 2					
For Intercept					
Intercept	0.491	0.050	9.765	25	<0.001
EF	−0.057	0.030	−1.874	25	0.072
PTA	−0.000	0.001	−0.483	25	0.633
For Slope					
Fidelity index	0.975	0.113	8.637	25	<0.001
EF	0.021	0.067	0.313	25	0.757
PTA	0.003	0.003	1.109	25	0.279
STEP 3					
For Intercept					
Intercept	0.273	0.087	3.145	24	0.005
EF	−0.062	0.029	−2.160	24	0.040
PTA	−0.0004	0.001	−0.430	24	0.670
Age	0.003	0.001	2.388	24	0.025
For Slope					
Fidelity index	1.081	0.280	3.867	24	0.001
EF	0.024	0.069	0.344	24	0.733
PTA	0.003	0.003	1.076	24	0.293
Age	−0.001	0.003	−0.433	24	0.669

Amount of hearing loss (PTA) is average of thresholds at 1, 2, 3, and 4 kHz in the test ear.

TABLE 7 | Summary of hierarchical linear model for quality with processing speed (PS) as a predictor.

Fixed effect	Coeff.	Std. error	T-ratio	*d.f*	*P*-value
STEP 1					
For Intercept					
Intercept	0.412	0.020	20.783	26	<0.001
PS	0.059	0.031	1.905	26	0.067
For Slope					
Fidelity index	1.119	0.038	29.827	26	<0.001
PS	−0.047	0.068	−0.629	26	0.495
STEP 2					
For Intercept					
Intercept	0.428	0.040	10.786		<0.001
PS	0.057	0.031	1.872		0.072
PTA	−0.000	0.001	−0.432		0.669
For Slope					
Fidelity index	1.009	0.113	8.927	25	<0.001
PS	−0.038	0.067	−0.565	25	0.575
PTA	0.003	0.002	1.103	25	0.281
STEP 3					
For Intercept					
Intercept	0.239	0.095	2.525	24	0.019
PS	0.055	0.030	1.863	24	0.074
PTA	−0.000	0.001	−0.373	24	0.712
Age	0.003	0.001	1.968	24	0.060
For Slope					
Fidelity index	1.102	0.274	4.018	24	0.001
PS	−0.037	0.067	−0.544	24	0.591
PTA	0.003	0.002	1.079	24	0.292
Age	−0.001	0.003	−0.383	24	0.705

Amount of hearing loss (PTA) is average of thresholds at 1, 2, 3, and 4 kHz in the test ear.

working memory capacity may be at a disadvantage. The present results were in good agreement with that expectation. Specifically, listeners with low working memory capacity (as quantified by a RST) performed more poorly for a given amount of signal modification (as quantified by the envelope fidelity metric) compared to individuals with high working memory capacity. That difference occurred despite having similar amount of hearing loss and age. Our results were consistent with the literature in showing the effect of working memory capacity on speech recognition. They also add to the literature regarding single-feature manipulations, from fast-acting WDRC (e.g., Gatehouse et al., 2006) and FC (e.g., Arehart et al., 2013a).

We also hypothesized that listeners with low working memory capacity would be disproportionately affected by high amounts of signal modification. Results of HLM modeling of intelligibility slope supported this hypothesis, although the effect also depended on the amount of hearing loss. In a general sense, the statistical result highlights the accumulating factors, with the poorest recognition of distorted signals by listeners with more hearing loss and with low working memory capacity. Our data reinforce results of Neher (2014), in which substantial variance in intelligibility was explained by amount of hearing loss and by working memory capacity.

Speech quality ratings were related to signal fidelity, but not to working memory capacity. There was a small effect of executive function on quality. Although our measure relied on rated speech quality rather than preference, and although we used the addition of background noise rather than noise reduction, this is generally consistent with Neher's (2014) finding that the preferred noise reduction setting depended on executive function (assuming that sound quality is a criterion for preference).

From a diagnostic standpoint, it is of interest to know whether one cognitive factor (here, working memory capacity, executive function, or processing speed) is a stronger predictor of intelligibility in adverse listening conditions. We hypothesized that individuals with lower executive function and/or slower processing speed might be more affected when adverse listening environments are created by signal modification. However, processing speed and executive function did not explain a significant proportion of the variance in speech intelligibility. Neher (2014) also examined the influence of executive function (specifically, the ability to maintain focus on relevant information) on speech modified by hearing-aid (noise reduction) processing. Consistent with our results, Neher reported that executive function accounted for a very small portion (3%) of the variance in a speech intelligibility task, and reported weak correlations among working memory (via a RST) and executive function. Overall, these findings suggest minimal influence of processing speed and executive function on speech intelligibility, but some qualifications are worth noting. First, in the present data and in Neher

(2014), working memory capacity was measured using a linguistic paradigm, while processing speed and executive function were measured using non-linguistic paradigms. It is likely that these non-linguistic paradigms did not capture the variability in top-down linguistic processing of sentence stimuli, which is a critical ability exploited by older listeners to compensate for distorted speech signals in challenging listening situations (Pichora-Fuller, 2008). Second, the speech intelligibility tasks used in both studies were directed speech tasks, in the sense that the listener's attention was pre-focused on the speech-in-noise signal. That presentation differs from many everyday situations in which the listener must direct attention among different talkers, potentially engaging executive function to a greater extent. It is possible that other measures of executive function and/or other speech scenarios might produce different results.

The present data (following the recent paper by Neher, 2014) add a multi-dimensional understanding of the relationship between working memory capacity and the characteristics of the speech signal, demonstrating that the relationship persists when signal modification is introduced via a combination of signal processing approaches. From a research perspective, these data are important as we refine our understanding of the role of working memory in adverse situations. From a translational perspective, these findings provide support for the idea that individuals with low working memory capacity might achieve better intelligibility with signal processing that maintains the fidelity of the signal envelope. However, more study is needed to explore the boundaries of the effect with regard to speech materials, noise type, and other aspects of listening, before such recommendations can be implemented in clinical practice. In particular, other aspects of hearing aid processing may produce different results. For example, the goal of noise suppression is to restore changes to the speech envelope caused by additive noise. Therefore, the cumulative effects of hearing aid signal processing that combines noise suppression with fast-acting WDRC and FC may differ from the results reported here. Finally, in the present study, the signal processing parameters were selected relative to our experimental goals, rather than customized for individual listeners. In future work, it will be important to consider both the effects of combined signal processing and customization of that processing to listener needs.

Acknowledgments

The authors thank Peggy Nelson for sharing speech materials, Thomas Lunner for providing the reading span test, Akira Miyake for guidance regarding the executive function and processing speed measures, Ramesh Kumar Muralimanohar for support with software development, Laura Mathews for assistance with data collection, and Rosalinda Baca for statistical analysis. A portion of these data was presented at the 2014 International Hearing Aid Conference, Tahoe City, CA. This work was supported by the National Institutes of Health (grant R01 DC012289 to PS and KA) and by a grant to the University of Colorado by GN ReSound (KA, JK).

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47 (Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

Alexander, J. M., Kopun, J. G., and Stelmachowicz, P. G. (2014). Effects of frequency compression and frequency transposition on fricative and affricate perception in listeners with normal hearing and mild to moderate hearing loss. *Ear Hear.* 35, 519–532. doi: 10.1097/AUD.0000000000000040

Anderson, M. C., Arehart, K. H., and Kates, J. M. (2009). The acoustic and perceptual effects of series and parallel processing. *EURASIP J. Adv. Signal Process.* 2009:619805. doi: 10.1155/2009/619805

Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). Aging affects neural precision of speech encoding. *J. Neurosci.* 32, 14156–14164. doi: 10.1523/JNEUROSCI.2176-12.2012

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4357–4362. doi: 10.1073/pnas.1213555110

ANSI. (2009). *American National Standards Institute Specification of Hearing Aid Characteristics*. New York, NY: ANSI.

Arehart, K. H., Kates, J. M., and Anderson, M. C. (2010). Effects of noise, nonlinear processing, and linear filtering on perceived speech quality. *Ear Hear.* 31, 420–436. doi: 10.1097/AUD.0b013e3181d3d4f3

Arehart, K. H., Kates, J. M., Anderson, M. C., and Harvey, L. O. Jr. (2007). Effects of noise and distortion on speech quality judgments in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 122, 1150–1164. doi: 10.1121/1.2754061

Arehart, K. H., Souza, P., Baca, R., and Kates, J. M. (2013a). Working memory, age, and hearing loss: susceptibility to hearing aid distortion. *Ear Hear.* 34, 251–260. doi: 10.1097/AUD.0b013e318271aa5e

Arehart, K. H., Souza, P., Lunner, T., Syskind Pedersen, M., and Kates, J. M. (2013b). "Relationship between distortion and working memory for digital noise-reduction processing in hearing aids," in *Presented at the Acoustical Society of America Convention* (Montreal, QC). doi: 10.1121/1.4805834

Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* 4, 417–423. doi: 10.1016/S1364-6613(00)01538-2

Besser, J., Koelewijn, T., Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2013). How linguistic closure and verbal working memory relate to speech recognition in noise–a review. *Trends Amplif.* 17, 75–93. doi: 10.1177/1084713813495459

Byrne, D., Dillon, H., Ching, T., Katsch, R., and Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures. *J. Am. Acad. Audiol.* 12, 37–51.

Chung, K. (2007). Effective compression and noise reduction configurations for hearing protectors. *J. Acoust. Soc. Am.* 121, 1090–1101. doi: 10.1121/1.2409859

Cox, R. M., Alexander, G. C., and Gilmore, C. (1987). Development of the Connected Speech Test (CST). *Ear Hear.* 8, 119S–126S. doi: 10.1097/00003446-198710001-00010

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Davies-Venn, E., and Souza, P. (2014). The role of spectral resolution, working memory, and audibility in explaining variance in susceptibility to temporal envelope distortion. *J. Am. Acad. Audiol.* 25, 592–604. doi: 10.3766/jaaa.25.6.9

Dupuis, K., Marchuk, V., Pichora-Fuller, M. K., Chasteen, A. L., Singh, G., and Smith, S. L. (2013). "Sensory loss and performance on the Montreal Cognitive Assessment," in *Presented at the Aging and Speech Communication Conference* (Bloomington, IN).

Ellis, R. J., and Munro, K. J. (2015). Benefit from, and acclimitization to, frequency compression hearing aids in experienced adult hearing-aid users. *Int. J. Audiol.* 54, 37–47. doi: 10.3109/14992027.2014.948217

Eriksen, B. A., and Ericksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Foo, C., Rudner, M., Rönnberg, J., and Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *J. Am. Acad. Audiol.* 18, 618–631. doi: 10.3766/jaaa.18.7.8

Franck, B. A. M., van Kreveld-Box, C. S. G. M., and Dreschler, W. A. (1999). Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. *J. Acoust. Soc. Am.* 106, 1452–1464. doi: 10.1121/1.428055

Fry, A. F., and Hale, S. (1996). Processing speed, working memory, and fluid intelligence: evidence for a developmental cascade. *Psychol. Sci.* 7, 237–241. doi: 10.1111/j.1467-9280.1996.tb00366.x

Gabrielsson, A., Schenkman, B. N., and Hagerman, B. (1988). The effects of different frequency responses on sound quality judgments and speech intelligibility. *J. Speech Hear. Res.* 31, 166–177. doi: 10.1044/jshr.3102.166

Gatehouse, S., Naylor, G., and Elberling, C. (2006). Linear and nonlinear hearing aid fittings–2. Patterns of candidature. *Int. J. Audiol.* 45, 153–171. doi: 10.1080/14992020500429484

George, E. L., Zekveld, A. A., Kramer, S. E., Goverts, S. T., and Festen, J. M. (2007). Auditory and nonauditory factors affecting speech recognition. *J. Acoust. Soc. Am.* 121, 2362–2375. doi: 10.1121/1.2642072

Hodgson, M., Steininger, G., and Razavi, Z. (2007). Measurement and prediction of speech and noise levels and the Lombard effect in eating establishments. *J. Acoust. Soc. Am.* 121, 2023–2033. doi: 10.1121/1.2535571

Humes, L., and Floyd, S. S. (2005). Measures of working memory, sequence learning, and speech recognition in the elderly. *J. Speech Lang. Hear. Res.* 48, 224–235. doi: 10.1044/1092-4388(2005/016)

Jenstad, L. M., and Souza, P. E. (2005). Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *J. Speech Lang. Hear. Res.* 48, 651–667. doi: 10.1044/1092-4388(2005/045)

Kates, J. M. (2008). *Digital Hearing Aids*. San Diego, CA: Plural Publishing.

Kates, J. M., and Arehart, K. H. (2014). The Hearing Aid Speech Perception Index (HASPI). *Speech Commun.* 65, 75–93. doi: 10.1016/j.specom.2014.06.002

Kates, J. M., Arehart, K. H., and Souza, P. (2013). Integrating cognitive and peripheral factors in predicting hearing-aid processing benefit. *J. Acoust. Soc. Am.* 134, 4458–4469. doi: 10.1121/1.4824700

Lunner, T., and Sundewall-Thoren, E. (2007). Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J. Am. Acad. Audiol.* 18, 604–617. doi: 10.3766/jaaa.18.7.7

McAulay, R. J., and Quatieri, T. F. (1986). Speech analysis/synthesis based on a sinusoidal representation. *IEEE Trans. Acoust. Speech Signal Process.* 34, 744–754. doi: 10.1109/TASSP.1986.1164910

McDermott, H. J. (2011). A technical comparison of digital frequency-lowering algorithms available in two current hearing aids. *PLoS ONE* 6:e22358. doi: 10.1371/journal.pone.0022358

Miyake, A., Emerson, M. J., and Friedman, N. P. (2000). Assessment of executive functions in clinical settings: problems and recommendations. *Semin. Speech Lang.* 21, 169–183. doi: 10.1055/s-2000-7563

Moore, B. C. (2004). Dead regions in the cochlea: conceptual foundations, diagnosis and clinical applications. *Ear Hear.* 25, 98–116. doi: 10.1097/01.AUD.0000120359.49711.D7

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Neher, T. (2014). Relating hearing loss and executive functions to hearing aid users' preference for, and speech recognition with, different combinations of binaural noise reduction and microphone directionality. *Front. Neurosci.* 8:391. doi: 10.3389/fnins.2014.00391

Neher, T., Grimm, G., and Hohmann, V. (2014). Perceptual consequences of different signal changes due to binaural noise reduction: do hearing loss and working memory capacity play a role? *Ear Hear.* 35, e213–e227. doi: 10.1097/AUD.0000000000000054

Neher, T., Grimm, G., Hohmann, V., and Kollmeier, B. (2013). Do hearing loss and cognitive function modulate benefit from different binaural noise-reduction settings? *Ear Hear.* 35, e52–e62. doi: 10.1097/AUD.0000000000000003

Ohlenforst, B., Souza, P., and MacDonald, E. (2014). "Interaction of working memory, compressor speed and background noise characteristics," in *Presented at the American Auditory Society Conference*, (Scottsdale, AZ).

Olsen, W. O. (1998). Average speech levels and spectra in various speaking/listening conditions: a summary of the Pearson, Bennett & Fidell (1977) report. *Am. J. Audiol.* 7, 21–25. doi: 10.1044/1059-0889(1998/012)

Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *Int. J. Audiol.* 47, S72–S82. doi: 10.1080/14992020802307404

Pichora-Fuller, M. K., and Souza, P. (2003). Effects of aging on auditory processing of speech. *Int. J. Audiol.* 42, 2S11–2S16. doi: 10.3109/14992020309074638

Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods, 2nd Edn*. Thousand Oaks, CA: Sage Publishing.

Rönnberg, J., Arlinger, S., Lyxell, B., and Kinnefors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive function. *J. Speech Hear. Res.* 32, 725–735. doi: 10.1044/jshr.3204.725

Rönnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rosenthal, S. (1969). IEEE: recommended practices for speech quality measurements. *IEEE Trans. Audio Electroacoust.* 17, 227–246.

Rosetti, H. C., Lacritz, L. H., Cullum, C. M., and Weiner, M. F. (2011). Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology* 77, 1272–1275. doi: 10.1212/WNL.0b013e318230208a

Salthouse, T. A. (1991). Mediation of adult age difference in cognition by reductions in working memory and speed of processing. *Psychol. Sci.* 2, 179–183. doi: 10.1111/j.1467-9280.1991.tb00127.x

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biol. Psychol.* 54, 35–54. doi: 10.1016/S0301-0511(00)00052-1

Singer, J. D., and Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195152968.001.0001

Smeda, K., Wolters, F., and Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *J. Am. Acad. Audiol.* 26, 183–196. doi: 10.3766/jaaa.26.2.7

Souza, P., Arehart, K. H., Kates, J. M., Croghan, N. B., and Gehani, N. (2013). Exploring the limits of frequency lowering. *J. Speech Lang. Hear. Res.* 56, 1349–1363. doi: 10.1044/1092-4388(2013/12-0151)

Souza, P., and Sirow, L. (2014). Relating working memory to compression parameters in clinically fit hearing aids. *Am. J. Audiol.* 23, 394–401. doi: 10.1044/2014_AJA-14-0006

Zahorian, S. A., and Rothenberg, M. (1981). Principal-components analysis for low-redundancy encoding of speech spectra. *J. Acoust. Soc. Am.* 69, 832. doi: 10.1121/1.385539

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Souza, Arehart, Shen, Anderson and Kates. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 04 December 2014
doi: 10.3389/fnins.2014.00391

Relating hearing loss and executive functions to hearing aid users' preference for, and speech recognition with, different combinations of binaural noise reduction and microphone directionality

Tobias Neher*

Medical Physics and Cluster of Excellence Hearing4all, Oldenburg University, Germany

Edited by:
Mary Rudner, Linköping University, Sweden

Reviewed by:
Patrik Sörqvist, University of Gävle, Sweden
Pamela Souza, Northwestern University, USA

****Correspondence:***
Tobias Neher, Department of Medical Physics and Acoustics, Carl-von-Ossietzky University, D-26111 Oldenburg, Germany
e-mail: tobias.neher@uni-oldenburg.de

Knowledge of how executive functions relate to preferred hearing aid (HA) processing is sparse and seemingly inconsistent with related knowledge for speech recognition outcomes. This study thus aimed to find out if (1) performance on a measure of reading span (RS) is related to preferred binaural noise reduction (NR) strength, (2) similar relations exist for two different, non-verbal measures of executive function, (3) pure-tone average hearing loss (PTA), signal-to-noise ratio (SNR), and microphone directionality (DIR) also influence preferred NR strength, and (4) preference and speech recognition outcomes are similar. Sixty elderly HA users took part. Six HA conditions consisting of omnidirectional or cardioid microphones followed by inactive, moderate, or strong binaural NR as well as linear amplification were tested. Outcome was assessed at fixed SNRs using headphone simulations of a frontal target talker in a busy cafeteria. Analyses showed positive effects of active NR and DIR on preference, and negative and positive effects of, respectively, strong NR and DIR on speech recognition. Also, while moderate NR was the most preferred NR setting overall, preference for strong NR increased with SNR. No relation between RS and preference was found. However, larger PTA was related to weaker preference for inactive NR and stronger preference for strong NR for both microphone modes. Equivalent (but weaker) relations between worse performance on one non-verbal measure of executive function and the HA conditions without DIR were found. For speech recognition, there were relations between HA condition, PTA, and RS, but their pattern differed from that for preference. Altogether, these results indicate that, while moderate NR works well in general, a notable proportion of HA users prefer stronger NR. Furthermore, PTA and executive functions can account for some of the variability in preference for, and speech recognition with, different binaural NR and DIR settings.

Keywords: hearing loss, cognition, hearing aids, signal processing, individual differences, personalized treatment

INTRODUCTION

Substantial variability in outcome is a consistent finding in hearing aid (HA) research. This holds true for a broad range of HA technologies, including amplification (e.g., Gatehouse et al., 2006a,b), noise reduction (NR) processing (e.g., Lunner, 2003; Brons et al., 2013), microphone directionality (DIR; e.g., Ricketts and Mueller, 2000; Keidser et al., 2013), and frequency compression (e.g., Glista et al., 2009; Souza et al., 2013). Presumably, this variability is related to the fact that HA users can differ in terms of a multitude of peripheral, central-auditory, or cognitive characteristics, even if they have similar audiograms and ages (cf., CHABA, 1988). Consequently, it is of interest to identify associations between such user characteristics and HA users' response to different forms of HA processing, as this would enable the development of fitting rationales that can take these dependencies into account. This would then allow for more individualized HA fittings.

Generally speaking, however, knowledge of such associations is rather sparse. This holds true especially for HA technology other than amplification. What is more, findings from related research studies are not always easily reconcilable with each other. One case in point is the role that executive functions play for benefit from different types of HA processing. "Executive functions" is an umbrella term that is typically thought to encompass a diverse, but related and overlapping, set of cognitive abilities such as working memory, attention, inhibition, and mental flexibility (e.g., Chan et al., 2008). More recently, HA researchers have focused on how one of these abilities—working memory—impacts hearing-impaired listeners' response to different HA processing, including dynamic range compression, NR processing, and frequency compression. Taken together, these studies suggest that HA users with smaller working memory capacity fare better with less aggressive HA processing whereas HA users with larger working memory capacity fare better with more aggressive HA

processing (e.g., Lunner and Sundewall-Thorén, 2007; Arehart et al., 2013; Ng et al., 2013). In these studies, working memory capacity was typically assessed using a measure of reading span (RS) after Daneman and Carpenter (1980), while HA outcome was typically assessed using objective (e.g., speech recognition) measures.

In two previous studies, we also investigated the influence of RS on response to NR processing (Neher et al., 2013, 2014). In addition to RS, we controlled PTA by testing four age-matched groups of elderly hearing-impaired listeners exhibiting either smaller ("H+") or larger ("H−") PTA and either longer ("C+") or shorter ("C−") RS. In terms of HA processing, we used a binaural NR algorithm and varied its strength from inactive through moderate to strong. In terms of assessing outcome, we collected objective (e.g., speech recognition) and subjective (i.e., overall preference) data at fixed signal-to-noise ratios (SNRs) between −4 and 8 dB. For the objective outcomes, we found little evidence that RS and PTA modulate NR outcome. For overall preference, on the other hand, we found that C− listeners preferred strong over moderate NR despite poorer speech recognition due to greater speech distortion, whereas C+ listeners did not. These differences could indicate that C− listeners are more affected by noise than C+ listeners and therefore favor greater noise removal (even at the expense of added speech distortions), whereas C+ listeners prioritize fewer speech distortions.

The fact that we could only see a clear influence of RS in our preference data and that poorer RS was associated with preference for stronger NR is in contrast to the findings summarized above basically suggesting the opposite data pattern for objective HA outcomes. In view of this discrepancy and the general shortage of research dealing with relations between executive functions and subjective HA outcome, we wanted to scrutinize the influence of RS on preferred NR strength. In addition, we wanted to investigate the influence of PTA and input SNR. This aim was motivated by indications in our previous data (see Table 2 in Neher et al., 2013) that preference for strong NR increases with input SNR (mean preference for strong NR across listener groups: 40, 46, 51, and 57% at −4, 0, 4, and 8 dB SNR, respectively) and that H− listeners prefer stronger NR than H+ listeners (mean preference for inactive, moderate, and strong NR across SNRs: 5, 44, and 51% for H− listeners and 10, 44, and 46% for H+ listeners, respectively). Furthermore, we wanted to investigate the influence of preprocessing our stimuli with a directional microphone. Because it attenuates non-frontal signal components and thus their impact on the NR gains computed for, and applied to, the signal mixture, a forward-facing directional microphone can reduce the amount of distortion in a frontal speech signal (cf., Neher et al., 2014). Given that recent research has linked executive functions to susceptibility to distortion caused by HA processing (Lunner et al., 2009; Arehart et al., 2013), it is possible that less speech distortion due to directional preprocessing leads to stronger preference for strong NR, at least for HA users with certain cognitive profiles. Moreover, we wanted to determine if any associations between HA outcome and RS are also apparent for other measures of executive function. Previous research has shown that different measures of executive function are not necessarily strongly correlated (e.g., Gatehouse and Akeroyd, 2008; Neher et al., 2012), suggesting at least partially independent executive processes. It is therefore possible that different measures of executive function are related differently to HA outcome (e.g., that while shorter RS is related to greater benefit from less NR, poorer performance on another measure of executive function might be related to greater benefit from more NR). To address this possibility we included two additional measures of executive function. That is, we selected two visual measures that (1) were non-verbal in nature, (2) were designed to tap into other executive functions than the (verbal) RS measure, and (3) differed from each other in terms of the range of executive functions covered (broader vs. narrower). Our rationale for doing so was to find out if these relatively different measures would give rise to similar patterns of association with listeners' response to our HA conditions. Finally, to address the apparent discrepancy between objective and subjective HA outcomes alluded to above, we also measured speech intelligibility to find out if preference for, and speech recognition with, the different HA conditions are differentially related to PTA and executive functions.

In summary, the aims of the current study were to (1) replicate the previously observed association between RS and preferred binaural NR strength, (2) find out if the other measures of executive function give rise to similar patterns of association, (3) determine if PTA, input SNR, and DIR also modulate preferred NR strength, and (4) find out if for speech recognition results are similar. Due to the lack of comparable research, the current study was rather exploratory in nature. Nevertheless, based on the results summarized above we hypothesized that (1) poorer RS and larger PTA would be associated with stronger preference for strong NR, (2) RS and the other measures of executive function would be differentially related to preference for HA processing, (3) preference for strong NR would increase with input SNR, DIR would reduce the amount of speech distortion and thus potentially weaken any observed relations between preferred NR strength and the measures of executive function, and (4) the associations between speech recognition, PTA, and the measures of executive function would be different from those for preference.

MATERIALS AND METHODS

Ethical approval for all experimental procedures was obtained from the ethics committee of the University of Oldenburg.

PARTICIPANTS

Participants were recruited from a cohort of several hundred hearing-impaired listeners belonging to the database of the Hörzentrum Oldenburg, Germany. Selection criteria were bilateral sensorineural hearing losses, asymmetry in air-conduction thresholds of no more than 15 dB HL across ears for the standard audiometric frequencies from 0.5 to 4 kHz, and air-bone gaps of no larger than 15 dB HL at any audiometric frequency between 0.5 and 4 kHz. Furthermore, all participants were required to be habitual HA users with at least 9 months of HA experience, to have normal or corrected-to-normal vision according to the Snellen eye chart (i.e., 20/40 acuity or better), to have no history of any psychiatric disorders (e.g., depression), and to have a DemTect score of at least 9 (with a score of 8 being the cut-off point for suspected dementia; Kalbe et al., 2004). Initially, we

selected 120 participants who satisfied these criteria and administered the RS measure (see below) to them. For further testing, we then selected 60 participants whom we could stratify into four well-matched groups based on the medians of their PTA and RS data. This ("H±C±") approach was consistent with our previous studies except that we increased the sample size from 40 to 60 participants this time to allow us to investigate the effects of interest more fully. None of these participants had taken part previously. However, most of them had experience with similar research studies. Participants were paid on an hourly basis for their participation.

Table 1 summarizes the main characteristics for all 60 participants and the H+C+, H+C−, H−C+, and H−C− subgroups. Performing one-way analyses of variance (ANOVAs) with Bonferroni *post-hoc* analyses on the age, PTA, and RS data of these subgroups confirmed (1) the lack of significant differences in terms of age [$F_{(3,\ 56)} = 0.4$; $p > 0.7$; $\eta_p^2 = 0.02$], (2) significant differences in terms of PTA [$F_{(3,\ 56)} = 40.0$; $p < 0.0001$, $\eta_p^2 = 0.68$] between all pairs of subgroups with different hearing status (all $p < 0.0001$) but no significant difference in terms of PTA between any two subgroups with the same hearing status (all $p = 1.0$), and (3) significant differences in terms of RS [$F_{(3,\ 56)} = 41.7$; $p < 0.0001$, $\eta_p^2 = 0.69$] between all pairs of subgroups with different RS status (all $p < 0.0001$) but no significant difference in terms of RS between any two subgroups with the same RS status (all $p = 1.0$). Compared to the cohort we had tested previously, these participants had slightly lower age (all subgroups), slightly better RS (all subgroups), and slightly smaller PTA (H− subgroups).

MEASURES OF EXECUTIVE FUNCTION

To assess executive function we used the RS measure after Daneman and Carpenter (1980) and two subtests from the commercially available, clinically validated "TAP-M" test battery (Zimmermann and Fimm, 2012). The TAP-M test battery was developed to assess elderly persons in terms of fitness for driving. The two measures used here were the so-called "distractibility" (DIS) and "executive control" (EC) subtests.

Reading span (RS) measure

The RS measure is a visual, verbal measure of working memory capacity that is rather widely used in audiological research (e.g., Neher et al., 2011, 2013; Arehart et al., 2013; Desjardins and Doherty, 2013; Ng et al., 2013). Our implementation, which is based on psycholinguistically controlled test items, closely mimics that of other researchers (cf., Carroll et al., 2014). It consists of a training round comprising three trials (which we carried out as often as needed until the participant had understood the task) and a test round comprising 54 trials (which we carried out once). On each trial, short sentence segments are displayed on a screen one at a time at a rate of one word per 0.8 s. After three segments, there is a pause of 1.75 s, during which the participant has to respond either "yes" if the previous three segments made up a semantically correct sentence (e.g., "Das Mädchen–sang–ein Lied"; "The girl–sang–a song") or "no" if the previous three segments made up a semantically absurd sentence (e.g., "Die Flasche–trank–Wasser"; "The bottle–drank–water"). Following a sequence of sentences (three, four, five, or six, in random order), the participant is asked to recall either the first or final words of all the three, four, five, or six previous sentences in any order. As before, we used the percentage of correctly recalled first and final words presented across the 54 trials to assess performance.

Table 1 | Means (and ranges) for the age, PTA, RS, and EC_{PC} data of all 60 participants as well as the H+C+, H+C−, H−C+, and H−C− subgroups ($N = 15$ per subgroup).

	Age (yr)	PTA (dB HL)	RS (%-correct)	EC_{PC} (%-correct)
All participants	72 (60–82)	45 (28–67)	38 (19–57)	86 (35–100)
H+C+	73 (61–81)	36 (28–45)	46 (39–56)	88 (35–100)
H−C+	71 (64–82)	54 (46–67)	47 (39–57)	76 (53–100)
H+C−	72 (63–80)	38 (29–45)	30 (19–37)	90 (73–100)
H−C−	73 (60–81)	52 (46–62)	30 (22–35)	90 (78–100)

Distractibility (DIS) measure

The DIS subtest from the TAP-M test battery is a visual, non-verbal measure of executive function, which according to its developers taps into selective attention and inhibition (Zimmermann and Fimm, 2012). In the middle of a computer screen, happy or sad smiley symbols are presented for short instances of time. The participant has to respond as quickly as possible by pressing a button whenever a sad smiley appears, but not when a happy smiley appears. At irregular timing intervals, distractor stimuli (i.e., abstract shapes or symbols) appear somewhere near the middle of the screen. These distractors are colored to make them perceptually more salient than the smileys, which are shown in black and white only. The DIS measure consists of a training round comprising 11 trials (which we carried out as often as needed until the participant had understood the task) and a test round comprising 150 trials (which we carried out once). In the test round, 60 target smileys are presented, 30 of which are preceded by a distractor. On average, the (randomized) duration of a trial is 2.3 s. The distractor and target stimuli are separated in time by 0.5 s. Distractors remain on the screen for 1.5 s, while target stimuli are only visible for 0.15 s. In accordance with recommendations given in the TAP-M manual we decided to explore two DIS performance measures: (1) the difference in median response time between correctly responded to target stimuli with and without preceding distractors ("$DIS_{\Delta RT}$"), and (2) the difference in the proportion of correct responses (calculated by subtracting the number of missed targets and wrong responses from 30 and dividing the result by 30) between trials with and without preceding distractors.

Executive control (EC) measure

The EC subtest from the TAP-M test battery is a visual, non-verbal measure of executive function, which according to its developers taps into working memory, mental flexibility, selective attention, and inhibition (Zimmermann and Fimm, 2012). In the middle of a computer screen, red or blue numbers and letters are presented one at a time for 0.5 s. The participant has to respond as quickly as possible to red numbers by pressing a left button and to blue

letters by pressing a right button, and to ignore blue numbers and red letters. The EC measure consists of a training round comprising 10 trials with five target stimuli (which we carried out as often as needed until the participant had understood the task) and a test round comprising 80 trials with 40 target stimuli (which we carried out once). The timing interval between consecutive stimuli varies randomly between 2 and 3 s. In accordance with recommendations given in the TAP-M manual we decided to explore two EC performance measures: (1) the median response time to correctly responded to target stimuli ("EC_{RT}"), and (2) the proportion of correct responses calculated by subtracting the number of missed targets and wrong responses from 40 and dividing the result by 40 ("EC_{PC}"). Despite several training rounds one participant was unable to carry out this test successfully, so we abandoned it in his case.

PHYSICAL TEST SETUP

The auditory tests were carried out in a soundproof booth. Inside the booth two computer screens were located. One screen was used for displaying information to the participants. The other screen, which the participants were unable to see, was used by an experimenter for scoring the participants' responses during the speech recognition measurements. All test software was implemented in MatLab (MathWorks, Natick, USA). Audio playback was via an Auritec (Hamburg, Germany) Earbox Highpower soundcard and a pair of Sennheiser (Wennebostel, Germany) HDA200 headphones. Calibration was carried out using a Brüel & Kjær (B&K; Nærum, Denmark) 4153 artificial ear, a B&K 4134 1/2″ microphone, a B&K 2669 preamplifier, and a B&K 2610 measurement amplifier.

The RS, DIS, and EC measures were administered in a quiet well-lit room. A computer screen displaying the stimuli was positioned about 0.5 m in front of the participants' face. During the DIS and EC measurements, participants responded to the stimuli using two large hardware buttons supplied with the TAP-M test battery.

SPEECH STIMULI

The speech stimuli closely resembled those from our previous studies. They were based on recordings from the Oldenburg sentence material (Wagener et al., 1999), which consists of 120 sentences that are low in semantic context and that all follow the form "name verb numeral adjective object" (e.g., "Thomas has two large flowers"). To simulate a realistic complex listening situation we convolved the sentence recordings with pairs of head-related impulse responses (HRIRs). These HRIRs were measured in a large, reverberant cafeteria using a B&K head-and-torso simulator (HATS) equipped with two three-microphone behind-the-ear Siemens Acuris HA "dummies" (Kayser et al., 2009). Each dummy consisted of the microphone array housed in its original casing, but without any of the integrated amplifiers, speakers, or signal processors commonly used in HAs. For the purpose of the current study, we used HRIRs measured with the front and rear (but not the mid) microphones and a frontal source at a distance of 1 m from, and at the same height as, the HATS. Following convolution with these HRIRs, the speech signals ranged in length from 2.2 to 3.2 s. For the interfering signal, we used a recording made in the same cafeteria with the same setup during a busy lunch hour. On each trial, a 5-s extract from this recording was randomly chosen and processed to have 50-ms raised-cosine on- and offset ramps. The resultant signal was presented at a nominal sound pressure level of 65 dB. It was mixed with a given target sentence, which started 1.25 s after the cafeteria noise and which was adjusted in level to produce a given SNR.

HEARING AID PROCESSING

All signal processing was implemented on the Master Hearing Aid (MHA) research platform of Grimm et al. (2006). It included DIR, binaural NR, linear amplification, and headphone equalization and was carried out at a sampling rate of 16 kHz. Before presentation, stimuli were resampled to 44.1 kHz. A total of six HA conditions were tested, which we will refer to as $DIR_{off}NR_{off}$, $DIR_{off}NR_{mod}$, $DIR_{off}NR_{str}$, $DIR_{on}NR_{off}$, $DIR_{on}NR_{mod}$, and $DIR_{on}NR_{str}$. These conditions differed in terms of whether (1) pairs of omnidirectional ("DIR_{off}") or cardioid ("DIR_{on}") microphones were used and (2) the binaural NR scheme was set to inactive ("NR_{off}"), moderate ("NR_{mod}"), or strong ("NR_{str}") processing.

Microphone directionality (DIR)

To simulate a pair of omnidirectional microphones we used the speech and noise signals obtained through convolution with the HRIRs measured with the front microphones of the two HA dummies. To simulate two directional microphones we employed the speech and noise signals obtained through convolution with the HRIRs measured with the front and rear microphones of the two HA dummies. Using a simple delay-and-sum beamformer algorithm (Elko and Pong, 1995), we then processed the two microphone signals per HA dummy in such a way that we obtained a pair of static forward-facing cardioid microphones. To compensate for the high-pass characteristic that is typical of directional microphones (e.g., Dillon, 2012), we applied a 1024th-order finite impulse response (FIR) filter to the output of each cardioid microphone. This filter ensured that the cardioid microphones were matched in terms of frequency response to their omnidirectional counterparts for the frontal (0° azimuth) source direction. We then also applied a two-channel 1024th-order FIR filter to the left and right outputs of each pair of omnidirectional or cardioid microphones. This filter ensured that the pairs of omnidirectional and cardioid microphones were matched in terms of their interaural phase and level differences for the frontal (0° azimuth) source direction. Directional microphone arrays are very sensitive to inter-microphone mismatch, which can result in considerable distortion of interaural cues (Van Den Bogaert et al., 2005). Thus, by post-processing the microphone signals in this manner, we made sure that the frontal target signals of our stimuli sounded highly similar across the omnidirectional and cardioid settings.

Binaural noise reduction (NR)

The binaural NR scheme was identical to that from our previous study (see Neher et al., 2013 for details). In short, it consisted of a Fast Fourier Transform-based filterbank with 12 frequency bands covering an 8-kHz bandwidth. Using a 40-ms integration

time constant, the binaural coherence (or interaural similarity) of the left and right input signals is first estimated in each frequency band. These estimates can take on values between 0 and 1. A value of 0 corresponds to fully incoherent (or diffuse) sound, while a value of 1 corresponds to fully coherent (or directional) sound. Because of diffraction effects around the head, the binaural coherence is always high below about 1 kHz. At higher frequencies, the coherence is low for diffuse and reverberant signal components, but high for the direct sound from nearby sources. Due to the spectro-temporal fluctuations contained in speech, the ratio between (undesired) incoherent and (desired) coherent signal components may vary across time and frequency. By applying appropriate time- and frequency-dependent gains this ratio can be improved. These gains are derived by applying an exponent, α, to the coherence estimates. As in our previous study, we tested three values of α: 0, 0.75, and 2. In this manner, we could vary the NR strength from inactive ($\alpha = 0$) through moderate ($\alpha = 0.75$) to strong ($\alpha = 2$).

Linear amplification and headphone equalization

To ensure adequate audibility we spectrally shaped all speech stimuli according to the National Acoustic Laboratories-Revised Profound (NAL-RP) prescription rule (Byrne et al., 1991). Specifically, for each participant we determined the required gain at 250, 500, 1000, 1500, 2000, 3000, 4000, and 6000 Hz and mapped the resultant values onto the MHA filterbank using interpolation techniques. Finally, we processed the left and right channels of each stimulus with a 32nd-order FIR filter that compensated for the uneven magnitude response of the headphones.

Physical effects

The chosen HA conditions gave rise to a number of physical effects, which are illustrated in **Figure 1** for one channel of an example stimulus with an input SNR of 4 dB. The panels on the left-hand side show, for each HA condition, the waveforms of the speech and noise signals at the output of the simulated HA. The panels on the right-hand side show the spectrograms of the signal mixtures. The dominant effect of moderate and especially strong NR is to suppress incoherent signal components above about 1 kHz. To quantify the physical effects of our HA conditions we calculated the speech-weighted SNR improvement ("ΔAI-SNR") for input SNRs of −4, 0, and 4 dB using a 2-min speech-in-noise stimulus. That is, we first estimated the SNR improvement relative to $DIR_{off}NR_{off}$ in one-third octave bands and then took the scalar product of these estimates and the one-third octave band importance function from the Speech Intelligibility Index (ANSI, 1997). **Table 2** summarizes the results. Relative to the omnidirectional setting, the cardioid setting led to a ΔAI-SNR of 3.3 dB, irrespective of input SNR. Furthermore, ΔAI-SNR increased with NR strength (e.g., from 1.7 dB for $DIR_{off}NR_{mod}$ to 2.8 dB for $DIR_{off}NR_{str}$ at 0 dB SNR) and input SNR (e.g., from

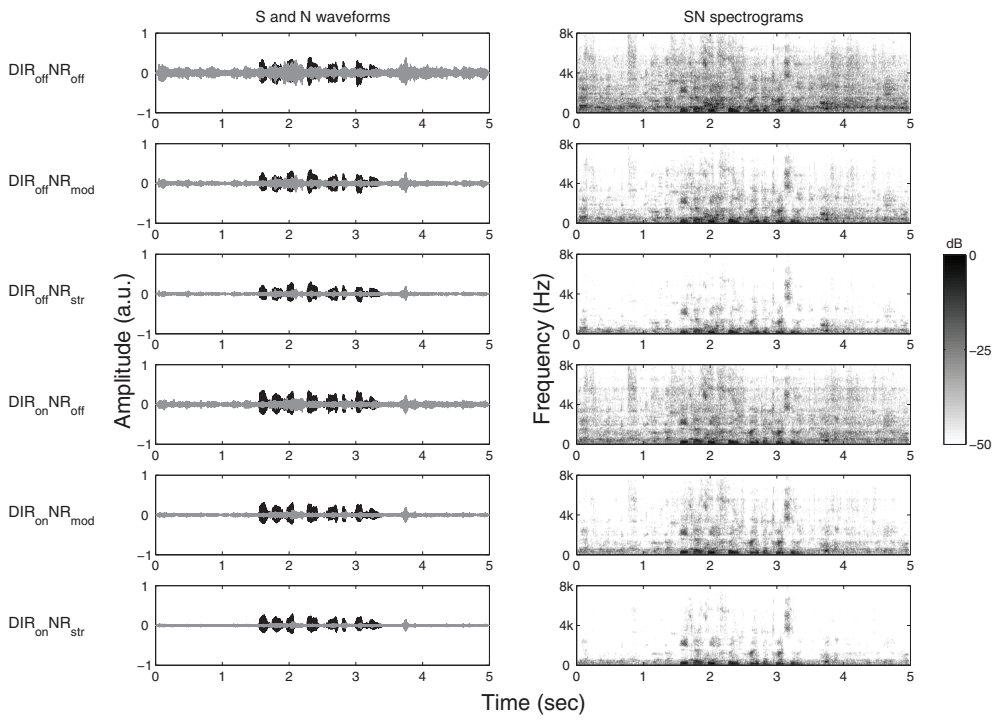


FIGURE 1 | Graphical illustration of the effects of $DIR_{off}NR_{off}$, $DIR_{off}NR_{mod}$, $DIR_{off}NR_{str}$, $DIR_{on}NR_{off}$, $DIR_{on}NR_{mod}$, and $DIR_{on}NR_{str}$ processing on (one channel of) an example stimulus with an input SNR of 4 dB. Panels on the left-hand side show time waveforms of the target speech, S (black) and the cafeteria noise, N (gray). Panels on the right-hand side show corresponding spectrograms for the signal mixtures, SN. a.u. denotes arbitrary units.

Table 2 | Speech-weighted SNR improvement (ΔAI-SNR) relative to $DIR_{off}NR_{off}$ for $DIR_{off}NR_{mod}$, $DIR_{off}NR_{str}$, $DIR_{on}NR_{off}$, $DIR_{on}NR_{mod}$, and $DIR_{on}NR_{str}$ and input SNRs of −4, 0, and 4 dB.

HA condition	ΔAI-SNR (dB)		
	−4 dB SNR	0 dB SNR	4 dB SNR
$DIR_{off}NR_{mod}$	1.0	1.7	2.3
$DIR_{off}NR_{str}$	1.5	2.8	3.8
$DIR_{on}NR_{off}$	3.3	3.3	3.3
$DIR_{on}NR_{mod}$	4.6	5.1	5.5
$DIR_{on}NR_{str}$	5.4	6.4	7.0

1.5 dB at −4 dB SNR to 3.8 dB at 4 dB SNR for $DIR_{off}NR_{str}$). It is also worth noting that, with the cardioid setting, the ΔAI-SNRs brought about by moderate and strong NR increased by, respectively, 0.3 and 0.6 dB at −4 dB SNR and by, respectively, 0.2 and 0.3 dB at 0 dB SNR; at 4 dB SNR, microphone mode basically had no influence on the ΔAI-SNRs due to moderate and strong NR.

In addition to SNR improvement, we quantified the amount of speech distortion caused by our HA conditions. To that end, we analyzed the stimuli from the ΔAI-SNR calculations using the Hearing Aid Speech Quality Index (HASQI; Kates and Arehart, 2014). HASQI assesses the amount of signal degradation in a processed stimulus relative to an unprocessed reference stimulus. It returns a value between 0 and 1, with 0 indicating very low fidelity and 1 indicating perfect fidelity. Because we were interested in the adverse effects of NR, we used the inactive NR setting as reference for the moderate and strong NR settings. Because we were also interested in the effects of directional preprocessing we performed these analyses separately for the omnidirectional and cardioid settings. In each case, we analyzed the target speech signals processed with the NR gains computed for the corresponding signal mixtures.

The HASQI values that we obtained ranged from 0.59 for strong NR without directional preprocessing at −4 dB SNR to 0.88 for moderate NR with directional preprocessing at 4 dB SNR (see **Table 3**). As expected, signal fidelity increased with SNR (mean HASQI values across NR and DIR setting: 0.72, 0.75, and 0.78 for −4, 0, and 4 dB SNR, respectively) and decreased with NR strength (mean HASQI values across SNR and DIR setting: 0.85 and 0.65 for moderate and strong NR, respectively). Furthermore, directional preprocessing had a positive effect on signal fidelity (mean HASQI values across SNR and NR setting: 0.74 and 0.76 for DIR_{off} and DIR_{on}, respectively). Altogether, these data show that the efficacy of our NR scheme increased with SNR, in terms of both SNR improvement and speech quality. Furthermore, not only did the cardioid setting lead to a considerable SNR improvement, it also reduced the speech distortion caused by moderate and strong NR.

Table 3 | Speech distortion (as measured using HASQI) caused by moderate and strong NR for the omnidirectional (DIR_{off}) and cardioid (DIR_{on}) settings and input SNRs of −4, 0, and 4 dB.

NR strength	Speech distortion (HASQI)					
	−4 dB SNR		0 dB SNR		4 dB SNR	
	DIR_{off}	DIR_{on}	DIR_{off}	DIR_{on}	DIR_{off}	DIR_{on}
Moderate	0.82	0.85	0.84	0.86	0.86	0.88
Strong	0.59	0.63	0.64	0.67	0.68	0.71

SPEECH RECOGNITION MEASUREMENTS

Consistent with our earlier studies, we determined speech recognition at −4 and 0 dB SNR. Since we had previously observed good test-retest reliability for similar measurements at these SNRs, we only made one measurement per condition. For the current study, we distributed the 12 measurements (6 HA conditions × 2 SNRs) in such a way that, at each of the two visits per participant (see Test protocol), three measurements per SNR were performed, each of the six HA conditions was tested once, and that the order of presentation was randomized. Furthermore, we started each visit with two training measurements carried out with $DIR_{on}NR_{off}$ processing at 4 and then 0 dB SNR. In total, each participant therefore completed 16 measurements. For each of these, we used a different test list (consisting of 20 five-word sentences each) and also balanced the lists across participants. Following the presentation of a stimulus, participants had to repeat the words they had understood, which an experimenter scored using a graphical user interface (GUI).

OVERALL PREFERENCE JUDGMENTS

For the preference judgments, we asked our participants to imagine being inside the cafeteria and wanting to communicate with the speaker of the sentences. They then had to compare a given pair of HA settings and decide which one they preferred overall. In doing so, they were instructed to pay attention to both target speech and background noise. Test conditions were identical to the speech recognition measurements, except that we also tested at 4 dB SNR. On each trial, six 5-s stimuli were generated as described above and concatenated, resulting in a 30-s stimulus. Comparisons were blocked by SNR. Different (randomly selected) speech signals and noise extracts were used for the different SNRs. Using a GUI and a touch screen, participants controlled playback of the (looped) stimuli and entered their responses. Participants completed four or five rounds of preference judgments (see Test protocol). One round consisted of 45 pairwise comparisons (3 SNRs × 15 possible combinations of the six HA conditions) in randomized order. At the start of the first round, six trials were initially presented for training purposes at 0 dB SNR. Presentation of the HA conditions was balanced in that the order of allocation of a given pair of HA conditions to the two buttons controlling playback was switched from one round to the next (e.g., $DIR_{off}NR_{off}$ vs. $DIR_{off}NR_{str}$ in the first round and $DIR_{off}NR_{str}$ vs. $DIR_{off}NR_{off}$ in the second round). The different rounds were not exact retests, as all stimuli were newly generated at the beginning of a round.

TEST PROTOCOL

All participants attended two 1.5-h visits. Each visit started with the speech recognition measurements (ca. 25 min) followed by the preference judgments. At the first visit, each participant

completed two rounds of preference judgments (ca. 20 min each). At the second visit, 35 participants completed another two rounds of preference judgments, while the other participants were able to complete three rounds each within the allotted time. After the speech recognition measurements and in-between the preference judgments participants were asked to take 5-min breaks.

STATISTICAL ANALYSES

In preparation for the statistical analyses, we divided the speech scores by 100 and transformed them into rationalized arcsine units (RAU; Studebaker, 1985). Furthermore, we converted the preference judgments into scores ranging from 0 to 1 by calculating, for each SNR, the total number of times a given HA condition was preferred to the other five conditions and then dividing the result by the total number of comparisons per condition (e.g., David, 1963; Arehart et al., 2007; Anderson et al., 2009). To avoid the influence of extreme values on our results and to normalize the variance in our datasets we excluded scores more than three times the interquartile range away from the lower and upper quartiles of a given dataset. Thus, we removed the DIS data of two participants. Furthermore, we excluded one participant altogether as, despite belonging to the H+C+ group, her speech scores were extraordinarily poor (grand average speech recognition: 11% correct). Finally, we also arcsine-transformed the proportions of correct responses from the EC measure.

Next, we carried out regression analyses with the aim of identifying the most predictive sets of between-subject factors for the speech and preference scores. Consistent with the H±C± approach we had used previously, we dichotomized the chosen predictors using a median split. In this way, we obtained two subgroups (or mean scores) per predictor, one denoting better ability (e.g., smaller PTA or longer RS) and one denoting worse ability (e.g., larger PTA or shorter RS). In a few cases, individual scores were equivalent to the overall median of a given dataset and thus had to be excluded. Subgroups therefore differed in size, but in no case included fewer than 24 individual scores. To test for statistically significant differences among our experimental variables we then performed mixed-model ANOVAs. Whenever appropriate, we corrected for violations of sphericity using the Greenhouse-Geisser correction. Furthermore, we included age as a covariate in each model. To leave the within-subject factor sum of squares unaltered we first centered the age variable by subtracting the overall sample mean from each data point (cf., Fidell and Tabachnick, 2006; Van Breukelen and Van Dijk, 2007).

Because of differences in the way we measured speech recognition and in the way we analyzed the preference data between our previous and the current study, we did not have estimates of test-retest reliability available and thus could not perform any power analyses.

RESULTS

ANALYSIS AND SELECTION OF BETWEEN-SUBJECT FACTORS

To identify the most effective predictors for the speech and preference scores we performed a series of multivariate multiple regression analyses. Using age as our baseline model, we assessed the predictive power of PTA and the measures of executive function both separately and in different combinations. In this manner, we could determine the unique variance explained by each predictor as well as the total variance explained by a given set of predictors. For each model tested, we averaged the explained variance across the various datasets per outcome (speech recognition: 2 SNRs × 6 HA conditions = 12 datasets; overall preference: 3 SNRs × 6 HA conditions = 18 datasets) to determine its total predictive power.

For the speech scores, we found that age accounted for 8.1% of the variance, while of the remaining predictors PTA, RS, $DIS_{\Delta PC}$, and $DIS_{\Delta RT}$ were most effective, accounting for 28.2, 13.5, 12, and 11%, respectively (together with age). The most effective combination consisted of PTA, RS, and $DIS_{\Delta RT}$ (unique R^2: 20.1, 5.4, and 3.1%, respectively). Together with age, they accounted for 36.7% of the total variance in the speech scores (range across datasets: 30–46%).

For the preference scores, we found that age accounted for 3.5% of the variance, while of the remaining predictors PTA, EC_{RT}, EC_{PC}, and RS were most effective, accounting for 9.7, 6.2, 6.1, and 4%, respectively (together with age). The most effective combination consisted of PTA, EC_{PC}, and EC_{RT} (unique R^2: 6.1, 2.7, and 2.4%, respectively). Together with age, they accounted for 14.6% of the total variance in the preference scores. Closer inspection revealed that explained variance varied markedly across the 18 datasets (range: 1–27%). Predictive power was lower at −4 dB SNR (mean R^2: 8%) than at 0 and 4 dB SNR (mean R^2: 17 and 18%, respectively). Predictive power was also lower for the measurements made with moderate NR (mean R^2: 9%) than for those made with inactive and strong NR (mean R^2: 18 and 16%, respectively). For the measurements made with the omnidirectional and cardioid settings predictive power was similar (mean R^2: 16 and 13%, respectively). It is also worth noting that, in contrast to our expectations, RS was an ineffective predictor of preferred HA condition. This will be discussed further below.

To complete the above analysis we computed pairwise Pearson's r correlation coefficients. The largest correlations that we found were the ones between RS and EC_{PC} and between PTA and EC_{PC}, which were both rather weak (both $r = 0.31$, $p = 0.02$).

SPEECH RECOGNITION

To further analyze the speech scores we performed an ANOVA with SNR and HA condition as within-subject factors, PTA, RS, and $DIS_{\Delta RT}$ as between-subject factors, and age as a covariate. Since we observed no statistically significant effects of $DIS_{\Delta RT}$ (i.e., the least predictive between-subject factor selected above) we removed it from the model. **Table 4** provides a summary of the results. The effects of PTA and RS were statistically significant, as were the effects of SNR, HA condition, and SNR × HA condition. Furthermore, PTA interacted with HA condition, while for RS the two-way interaction with SNR and HA condition was significant.

Figure 2 shows mean speech scores with 95% confidence intervals for the six HA conditions and two SNRs. As expected, speech recognition improved with SNR. To investigate the significant effect of HA condition further we carried out a series of planned contrasts. These revealed significant differences among all pairs of HA conditions (all $p < 0.05$) except for $DIR_{off}NR_{off}$ vs. $DIR_{off}NR_{mod}$ ($p > 0.1$). Thus, across the two SNRs moderate NR did not affect speech recognition when combined with

Table 4 | Results from the ANOVA performed on the speech scores.

Model term	*df*	*F*	*p*	η_p^2
BETWEEN-SUBJECT FACTORS/COVARIATES				
PTA	(1, 54)	14.7	<0.001	0.22
RS	(1, 54)	7.6	<0.01	0.12
Age	(1, 54)	8.7	<0.01	0.14
WITHIN-SUBJECT FACTORS				
SNR	(1, 54)	612	<0.00001	0.92
HA	(4.1, 221.1)	390	<0.00001	0.88
SNR × HA	(5, 270)	26.0	<0.00001	0.33
SNR × PTA	(1, 54)	6.4	<0.05	0.11
HA × PTA	(4.1, 221.1)	2.6	<0.05	0.05
SNR × HA × RS	(5, 270)	3.3	<0.01	0.06

HA denotes HA condition. Model terms not shown were not statistically significant.

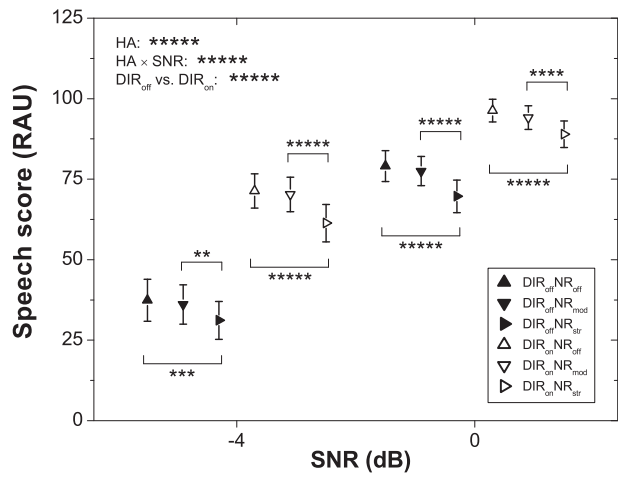


FIGURE 2 | Mean speech scores for the six HA conditions and two SNRs. Error bars show 95% confidence intervals. Horizontal bars denote significant differences (** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ***** $p < 0.00001$). HA denotes HA condition.

the omnidirectional setting, whereas in combination with the cardioid setting it led to a reduction of about 1.5 RAU ($p = 0.041$). Furthermore, strong NR reduced speech recognition by about 7 RAU across the two SNRs irrespective of microphone mode, while relative to no DIR the cardioid setting improved speech recognition by about 25 RAU averaged across SNR and NR setting.

Figure 3 shows the speech scores of the two PTA subgroups (left panel) and the two RS subgroups (right panel) for the different HA conditions. As expected, the "smaller PTA" and "better RS" subgroups achieved better speech recognition than the "larger PTA" and "worse RS" subgroups. To investigate the significant interaction between PTA and HA condition further we carried out series of planned contrasts on the data from the "larger PTA" and "smaller PTA" subgroups. For the "smaller PTA" subgroup, we found that the decrement in speech recognition due to strong (relative to inactive) NR was basically unaffected by the microphone setting (7.7 vs. 7.3 RAU), whereas for the "larger PTA" subgroup it was slightly larger with the cardioid setting (8.0 vs. 10.0 RAU). These results suggest that in terms of speech recognition HA users with larger PTA fare slightly worse with strong NR than HA users with smaller PTA if the NR is applied in conjunction with a pair of directional microphones.

To investigate the significant two-way interaction between SNR, HA condition, and RS further we carried out separate ANOVAs on the data from −4 and 0 dB. We found that RS interacted with HA condition at 0 dB ($p = 0.026$) but not at −4 dB ($p = 0.075$). Thus, we analyzed the 0 dB data further by carrying out series of planned contrasts on the data from the "better RS" and "worse RS" subgroups. For the "worse RS" subgroup, we found that the decrement in speech recognition due to strong (relative to inactive) NR was basically unaffected by microphone setting (7.6 vs. 7.0 RAU), whereas for the "better RS" subgroup it was slightly larger with the omnidirectional setting (11.2 vs. 7.6 RAU). These results suggest that in terms of speech recognition HA users with larger RS fare slightly worse with strong NR than HA users with smaller RS if the NR is applied without directional microphones.

OVERALL PREFERENCE

Because the preference scores were proportional values reflecting how much a given HA condition was preferred to each of the other five HA conditions for a given SNR, we analyzed these scores further by performing three separate ANOVAs—one per SNR. In each model, we included HA condition as within-subject factor, PTA, EC_{PC}, and EC_{RT} as between-subject factors, and age as a covariate. Since we observed no statistically significant effects of EC_{RT} (i.e., the least predictive between-subject factor selected above) we removed it from the models. **Table 5** provides a summary of the results. For each SNR, we found a highly significant effect of HA condition. Furthermore, whereas neither PTA nor EC_{PC} interacted with HA condition at −4 dB, we found significant interactions between each of these factors and HA condition at 0 and 4 dB SNR. **Table 1** therefore also provides a summary of the EC_{PC} data.

Figure 4 shows the effect of HA condition on overall preference for each of the three SNRs tested. As already noted in the context of the regression analyses (see above), inter-individual variability in preferred NR strength was smallest for moderate NR and much larger for inactive and strong NR, especially at 0 and 4 dB SNR. To investigate the significant effect of HA condition further we carried out a series of planned contrasts on the data from each SNR. At −4 dB, we found that moderate NR was significantly preferred over inactive and strong NR with and without DIR (all $p < 0.00001$). Furthermore, we found that strong NR was significantly preferred over inactive NR without DIR ($p < 0.01$) but not over inactive NR with DIR ($p > 0.7$). At 0 dB, the pattern was very similar, although there was a tendency for preference for strong NR to increase, particularly so in combination with DIR. This trend continued at 4 dB such that moderate and strong NR were equally preferred both with and without DIR (both $p > 0.3$). In terms of directional benefit, we found a very strong preference for DIR over no DIR at all three SNRs (all $p < 0.00001$).

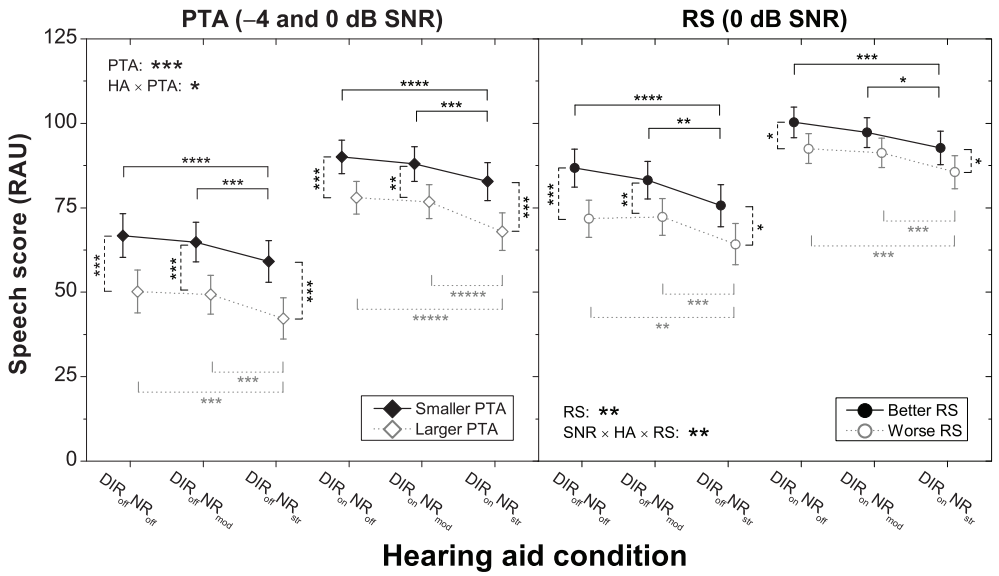


FIGURE 3 | Mean speech scores for the six HA conditions for the "smaller PTA" and "larger PTA" subgroups at −4 and 0 dB SNR (left panel) and the "better RS" and "worse RS" subgroups at 0 dB SNR (right panel). Error bars show 95% confidence intervals. Black solid horizontal bars denote significant differences for the "smaller PTA" and "better RS" subgroups, gray dotted horizontal bars denote significant differences for the "larger PTA" and "worse RS" subgroups, and black dashed vertical bars denote significant differences among pairs of subgroups for a given HA condition (*$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$, *****$p < 0.00001$). HA denotes HA condition.

Table 5 | Results from the ANOVAs performed on the preference scores corresponding to −4, 0, and 4 dB SNR.

Model term	−4 dB SNR				0 dB SNR				4 dB SNR			
	df	*F*	*p*	η_p^2	*df*	*F*	*p*	η_p^2	*df*	*F*	*p*	η_p^2
HA	(2.3, 108.2)	142	<0.00001	0.75	(1.8, 86.0)	71	<0.00001	0.60	(1.7, 81.0)	59	<0.00001	0.56
HA × PTA	(2.3, 108.2)	2.0	>0.1	0.04	(1.8, 86.0)	4.8	<0.05	0.09	(1.7, 81.0)	6.6	<0.01	0.12
HA × EC_{PC}	(2.3, 108.2)	1.2	>0.2	0.03	(1.8, 86.0)	3.8	<0.05	0.07	(1.7, 81.0)	3.4	<0.05	0.07

HA denotes HA condition. Model terms not shown were not statistically significant.

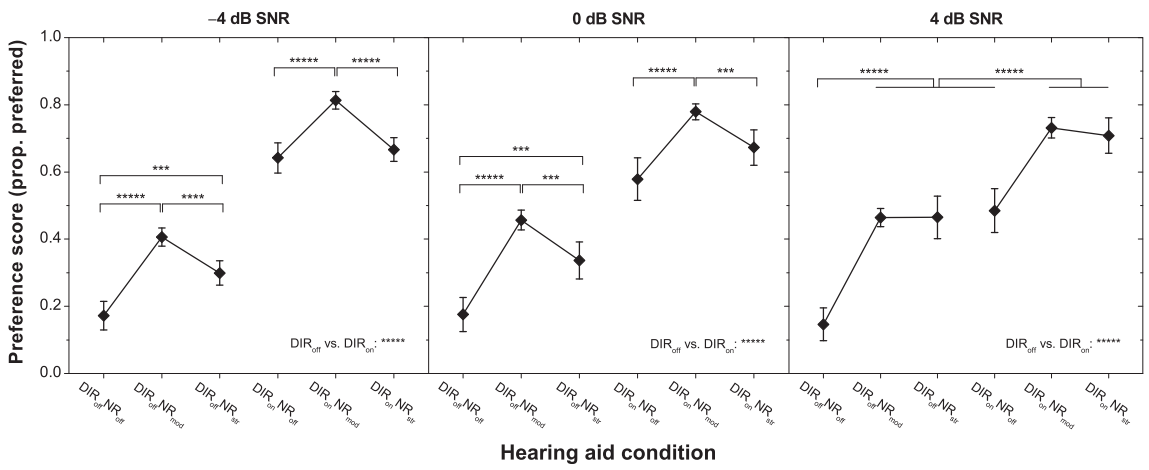


FIGURE 4 | Mean preference scores for the six HA conditions and SNRs of −4 dB (left panel), 0 dB (middle panel), and 4 dB (right panel). Error bars show 95% confidence intervals. Horizontal bars denote significant differences (***$p < 0.001$, ****$p < 0.0001$, *****$p < 0.00001$).

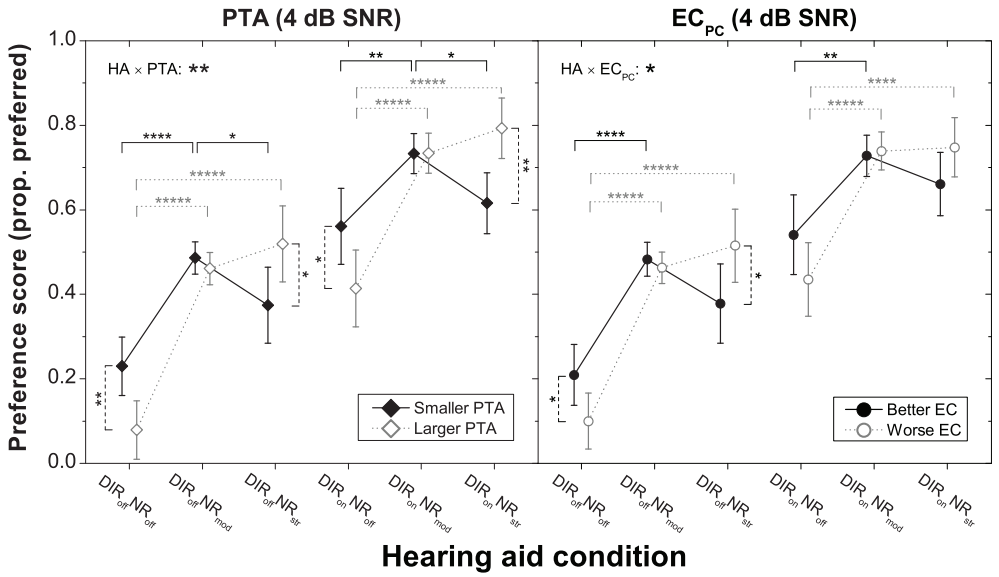


FIGURE 5 | Mean preference scores for the six HA conditions at 4 dB SNR for the "smaller PTA" and "larger PTA" subgroups (left panel) and the "better EC" and "worse EC" subgroups (right panel). Error bars show 95% confidence intervals. Black solid horizontal bars denote significant differences for the "smaller PTA" and "better EC" subgroups, gray dotted horizontal bars denote significant differences for the "larger PTA" and "worse EC" subgroups, and black dashed vertical bars denote significant differences among subgroups within a given HA condition (*$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$, *****$p < 0.00001$). HA denotes HA condition.

To scrutinize the significant interactions between HA condition, PTA, and EC_{PC} we carried out series of planned contrasts on the data from 0 to 4 dB SNR. Effects were clearest at 4 dB SNR and are therefore shown in **Figure 5**. For both microphone settings, the "larger PTA" subgroup more strongly disliked inactive NR than the "smaller PTA" subgroup, whereas for strong NR the situation was reversed (all $p < 0.05$). Similarly, for the omnidirectional (but not the cardioid) microphone setting the "worse EC" subgroup more strongly disliked inactive NR than the "better EC" subgroup, whereas for strong NR the situation was reversed (both $p < 0.05$). At 0 dB SNR, the picture was very similar although the differences in mean preference between the "smaller PTA" and "larger PTA" subgroups were no longer significant at the 5% level for the $DIR_{off}NR_{str}$ and $DIR_{on}NR_{off}$ conditions (both $p = 0.06$). The same was true for the difference in mean preference between the "better EC" and "worse EC" subgroups for the $DIR_{off}NR_{off}$ condition ($p = 0.08$). Finally, it should be noted that whereas at 0 dB SNR all subgroups preferred moderate NR the most, at 4 dB SNR the "larger PTA" and "worse EC" subgroups tended to more strongly prefer strong NR. Nevertheless, because of the considerable inter-individual variability in preference for strong NR, the mean scores for the moderate and strong NR settings did not differ statistically from each other (all $p > 0.1$).

Altogether, these results suggest that in terms of preference HA users with larger PTA fare better with stronger NR than HA users with smaller PTA irrespective of microphone mode. Similarly, they suggest that HA users with poorer EC performance also fare better with stronger NR than HA users with better EC performance, but only in combination with the omnidirectional setting.

DISCUSSION

The current study had four main aims: (1) to confirm the previously observed association between RS and preferred NR setting, (2) to find out if there are similar associations with the DIS and EC measures, (3) to investigate if PTA, input SNR, and DIR also modulate preferred NR setting, and (4) to find out if preference and speech recognition show similar relations to PTA and the measures of executive function. Regarding the first aim, we saw no indications in the data from the current study that RS interacts with preferred NR setting. Regarding the second aim, DIS did not affect preference for the various HA conditions either, whereas EC_{PC} could partly account for the observed inter-individual variability. Regarding the third aim, we found larger PTA to be associated with weaker preference for inactive NR and stronger preference for strong NR, preference for strong NR to increase with input SNR, and DIR to weaken the association between EC_{PC} and preferred NR setting. Regarding the fourth aim, we observed that PTA and the measures of executive function interacted differentially with preference and speech recognition. In the following sections, we discuss these results in more detail.

EFFECTS OF EXECUTIVE FUNCTIONS

As pointed out above, it is not uncommon to observe weak correlations among different measures of executive function, which was also the case in the current study (see Results). Presumably, this was at least partly because we had used *non-verbal* benchmarks for the *verbal* RS measure. We therefore had expected that these measures would give rise to different patterns of association with HA outcome, and this is also what we found.

In our previous study, listeners with shorter RS had preferred strong over moderate NR, whereas listeners with longer RS had

not (see Introduction). However, our current study revealed no influence of RS on preferred NR strength (nor on preferred microphone setting). For the current study, we had deliberately recruited new participants. One reason for the divergent results across studies concerning the influence of RS could therefore be that the salient characteristics were not sufficiently pronounced in the cohort tested this time—perhaps because we had screened potential candidates more rigorously. In fact, however, the RS scores of the cohorts from the previous and current study were very similar (mean RS scores: 38.2 vs. 36.0%-correct; coefficients of variation: 0.27 vs. 0.27), thereby ruling out such an explanation. Another reason for the conflicting results could be random sampling variation. In principle, it is possible that preference for NR strength is a very individual trait that is not easily captured by a given measure of executive function. If this were the case, it would not be possible to assess the influence of executive function on the NR strength preferred by elderly HA users reliably based on a few samples of that population.

Apart from RS, DIS was also unrelated to preference for HA condition. To recapitulate, we had included DIS as a non-verbal benchmark for the RS measure indexing different executive functions (i.e., selective attention and inhibition). Incidentally, the spread in the DIS data was notably larger (coefficient of variation = 1.6) than in the RS data. In spite of this, DIS failed to account for any of the inter-individual variability in our preference scores. We therefore conclude that this measure was not sensitive to the executive processes driving preference for the HA conditions tested here.

In contrast to the other measures of executive function, EC_{PC} was associated with preference for our HA conditions. This was despite the fact that the spread in the EC_{PC} data (coefficient of variation = 0.22) was smaller than in the DIS and RS datasets. Our motivation for including EC was to have another non-verbal benchmark for RS indexing a wider range of executive functions than DIS (i.e., working memory, mental flexibility, selective attention, and inhibition). At present, it is unclear why precisely EC_{PC} could explain some of the variability in our preference scores. We speculate that because of the relatively broad spectrum of executive functions it taps into it was in a better position to capture the executive processes governing our listeners' preference judgments. Future research should try to identify the precise factors driving the observed association, ideally with the help of a new cohort of HA users.

EFFECTS OF HEARING LOSS

Regarding hearing loss, our earlier study had indicated that listeners with larger PTA prefer stronger NR than listeners with smaller PTA (see Introduction), and the results from the current study were consistent with this.

Only a couple of studies seem to have investigated the influence of PTA on preferred NR setting so far. In one study, five single- or multichannel NR schemes were tested, including the binaural coherence-based algorithm tested by us (Luts et al., 2010). Groups of listeners with normal hearing (ages 16–52), flat hearing losses (ages 22–79), and sloping hearing losses (ages 51–80) participated. Outcome measures included speech recognition and overall preference. For most NR schemes, the changes in outcome were very similar across groups, suggesting a negligible influence of PTA. For the binaural coherence-based algorithm, however, a significant effect of listener group was observed. That is, whereas the two hearing-impaired groups preferred this type of NR over no processing, the normal-hearing listeners did not. In another study, Houben et al. (2012) investigated preferred NR strength for two single-channel algorithms. Ten normal-hearing listeners (ages 21–31) and seven listeners with sloping hearing losses in the mild to severe range (ages 25–61) participated. For both groups, considerable inter-individual differences in preferred NR strength were observed. Also, their data overlapped considerably, resulting in a non-significant group effect. However, due to the small sample size and the fact that no attempt was made to control for any other factors that may affect HA outcome (e.g., age or executive functions), this result is not particularly surprising.

In summary, the influence of PTA that we observed was consistent with our previous data and, broadly speaking, also the results of Luts et al. (2010). The fact that Luts et al. did not find a corresponding group difference for any of their other NR schemes raises the question of whether the observed influence of PTA only pertains to the binaural coherence-based algorithm tested here. This should be addressed by future research.

EFFECTS OF SNR AND MICROPHONE DIRECTIONALITY

Concerning the influence of SNR, our earlier study had indicated that preference for strong NR increases with input SNR (see Introduction), and the results from the current study were consistent with this. This dependency can be traced back to the fact that with higher input SNR the adverse effects of the NR processing (i.e., speech distortion) decreased while its positive effects (i.e., noise attenuation) increased, as confirmed by our technical analyses (see **Tables 2**, **3**). Consequently, the benefit from strong NR increasingly outweighed its unwanted side effects. Based on this interpretation, one would expect even stronger preference for strong NR above 4 dB SNR. In actual fact, this is what we observed in our previous study, as part of which we had also collected preference judgments at 8 dB SNR (see Introduction). Interestingly, we did not observe any effects of PTA or the measures of executive function at −4 dB SNR. Previously, we had observed rather poor reproducibility for NR preference ratings at −4 dB SNR, whereas at 0 and especially 4 dB SNR reproducibility had been much better (Neher et al., 2014). Perhaps because speech distortion was greatest at −4 dB SNR participants were unsure about their preferences, thereby leading to no consistent associations with PTA or the measures of executive function.

Concerning the influence of DIR, we observed a clear preference for the cardioid over the omnidirectional setting. This is consistent with the finding of other researchers that DIR is preferred when noise is present and the signal of interest is in front of, and relatively near to, the listener (e.g., Walden et al., 2004, 2005). Furthermore, we had hypothesized that because directional preprocessing can reduce the amount of speech distortion caused by NR this might affect the influence of executive functions on preferred NR setting. Our technical analyses confirmed an improvement in speech quality due to DIR (see **Table 3**). Our perceptual analyses revealed that the observed association between EC_{PC} and preferred NR strength only applied to the HA conditions without

DIR. Thus, these findings were consistent with our hypothesis. At first sight, they are also consistent with the idea that executive processes modulate susceptibility to HA distortion, as proposed by Lunner et al. (2009). According to their view, individual differences in working memory capacity determine listening success with specific types of HA technology. In particular, listeners with greater working memory capacity are thought to be better at segregating a target signal from any unwanted artifacts as they can deploy some of this capacity for explicit (as opposed to implicit or effortless) processing needed to match suboptimal input with phonologically based long-term representations in their mental lexicon (cf., Rönnberg, 2003; Rönnberg et al., 2008). Although this view is consistent with the results from a number of HA studies focusing on speech recognition outcomes (see Introduction), it seemingly disagrees with the effects apparent in our preference data. This is discussed further below.

OVERALL PREFERENCE vs. SPEECH RECOGNITION

In HA research, preference judgments and speech recognition measurements commonly produce divergent data patterns (e.g., Walden et al., 2005; Brons et al., 2013; Jensen et al., 2013). In view of this as well as our earlier results (see Introduction), we had expected PTA and the measures of executive function to be differentially related to our speech and preference scores. To summarize, our analyses of the preference scores had suggested that HA users with larger PTA fare *better* with stronger (i.e., more aggressive) NR, whereas for listeners with smaller PTA the opposite holds true (at 0 and 4 dB SNR with and without DIR). Furthermore, they had suggested that listeners with worse EC_{PC} performance fare also *better* with stronger NR, whereas for listeners with better EC_{PC} performance the opposite holds true (at 0 and 4 dB SNR without DIR). Our analyses of the speech scores, on the other hand, had suggested that HA users with larger PTA fare slightly *worse* with stronger NR than HA users with smaller PTA (at −4 and 0 dB SNR with DIR). Furthermore, they had suggested that HA users with worse RS performance fare slightly *better* with strong NR than HA users with better RS performance (at 0 dB SNR without DIR).

Taken together, there appears to be some consistency across our preference and speech recognition results concerning the influence of executive functions (but not PTA) on response to our HA conditions. Recall, however, that we used different measures of executive function for the analyses of the two datasets. This was because we had found the (linguistically based) RS measure to be predictive of the speech but not the preference scores, while for the (non-verbal) EC measure the opposite was true. Broadly speaking, this pattern of results is consistent with previous reports of the strongest associations between verbal measures of executive function (in particular RS) and speech recognition (cf., Akeroyd, 2008).

Importantly, the associations with EC_{PC} and RS that we found were in disagreement with the literature finding that HA users with longer RS fare better with more aggressive HA settings and vice versa (see Introduction). Incidentally, even though statistically significant, the across-subgroup effects of RS (and PTA) in our speech scores were on the order of a few percentage points only. One could speculate that for a clear influence of executive functions on the speech recognition with different HA conditions to emerge listeners need to be confronted with more pronounced signal distortions such as those caused by frequency compression (cf., Arehart et al., 2013). Some support for this is available from a recent study of Keidser et al. (2013) concerned with individual differences in speech recognition benefit from DIR—a type of HA technology that is typically free from any distortions for the target direction (e.g., Dillon, 2012)—which failed to find a clear influence of executive functions (and PTA). In principle, it is also possible that different executive functions interact differentially with the signal changes caused by different HA algorithms.

In summary, the reported influence of executive functions on response to HA signal processing differs somewhat across HA outcomes and studies. Future research in this field should therefore ideally focus on trying to reconcile the findings from different studies.

IMPLICATIONS FOR HEARING AID FITTING

The results from our study imply that moderate NR works well for the majority of elderly HA users, especially when applied in conjunction with DIR (see **Figures 2**, **4**). They also show that HA users experience benefit from NR processing at positive SNRs (see **Figure 4**) where at least some HA manufacturers curtail the efficacy of their NR schemes (cf. Smeds et al., 2010). Furthermore, our results suggest that a notable proportion of elderly HA users prefer strong over moderate NR. Because strong NR may interfere with speech intelligibility, it is important to be able to identify candidates for strong NR reliably. Although our analyses had revealed that PTA and EC_{PC} can partly account for the inter-individual variability in preference for inactive and strong NR, their predictive power was limited (with unique R^2 for PTA and EC_{PC} amounting to about 11 and 6.3%, respectively, at 0 and 4 dB SNR). In addition, mean preference scores for the various subgroups did not differ statistically across the moderate and strong NR settings (see **Figure 5**). A relevant question therefore is whether the combined predictive power of PTA and EC_{PC} is sufficiently large to allow determining candidature for moderate or strong NR. To address this we performed a supplementary ANOVA for which we grouped PTA and EC_{PC} into a single H±EC± factor (akin to the H±C± factor we had used previously). Results showed that the H–EC– subgroup significantly preferred $DIR_{off}NR_{str}$ over $DIR_{off}NR_{mod}$ at 4 dB SNR (mean preference scores: 0.55 vs. 0.44 scale points; $p = 0.029$). Otherwise no differences in preference for strong over moderate NR were observable (all $p > 0.16$).

Altogether, our results indicate the basic potential of individualizing NR based on PTA and (to a lesser extent) executive functions. Furthermore, they point toward a need for alternative diagnostic measures that can capture more of the variability in preference for different NR settings, and current work in our laboratory is concerned with this issue.

ACKNOWLEDGMENTS

The author thanks his colleagues at the Hörzentrum Oldenburg for their help with recruiting the participants and performing the measurements, Giso Grimm for support with the Master Hearing Aid, and Jim Kates for supplying the HASQI code. This

research was funded by the DFG Cluster of Excellence EXC 1077/1 "Hearing4all" and by Siemens Audiologische Technik, Erlangen, Germany. Parts of it were presented at the 2014 International Hearing Aid Research Conference, Lake Tahoe, California, Aug. 13–17.

REFERENCES

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, 53–71. doi: 10.1080/14992020802301142

Anderson, M. C., Arehart, K. H., and Kates, J. M. (2009). The acoustic and perceptual effects of series and parallel processing. *EURASIP J. Adv. Signal Process.* 2009:619805. doi: 10.1155/2009/619805

ANSI. (1997). *Methods for Calculation of the Speech Intelligibility Index (S3.5-1997).* New York, NY: American National Standards Institute.

Arehart, K. H., Kates, J. M., Anderson, M. C., and Harvey, L. O. Jr. (2007). Effects of noise and distortion on speech quality judgments in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* 122, 1150–1164. doi: 10.1121/1.2754061

Arehart, K. H., Souza, P., Baca, R., and Kates, J. M. (2013). Working memory, age, and hearing loss: susceptibility to hearing aid distortion. *Ear Hear.* 34, 251–260. doi: 10.1097/AUD.0b013e318271aa5e

Brons, I., Houben, R., and Dreschler, W. A. (2013). Perceptual effects of noise reduction with respect to personal preference, speech intelligibility, and listening effort. *Ear Hear.* 34, 29–41. doi: 10.1097/AUD.0b013e31825f299f

Byrne, D., Parkinson, A., and Newall, P. (1991). "Modified hearing aid selection procedures for severe/profound hearing losses," in *The Vanderbilt Hearing Aid Report II*, eds G. A. Studebaker, F. H. Bess, and L. B. Beck (Parkton, NC: York Press), 295–300.

Carroll, R., Meis, M., Schulte, M., Vormann, M., Kiessling, J., and Meister, H. (2014). Development of a German reading span test with dual task design for application in cognitive hearing research. *Int. J. Audiol.* doi: 10.3109/14992027.2014.952458. [Epub ahead of print].

CHABA. (1988). Speech understanding and aging. *J. Acoust. Soc. Am.* 83, 859–895. doi: 10.1121/1.395965

Chan, R. C., Shum, D., Toulopoulou, T., and Chen, E. Y. (2008). Assessment of executive functions: review of instruments and identification of critical issues. *Arch. Clin. Neuropsychol.* 23, 201–216. doi: 10.1016/j.acn.2007.08.010

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

David, H. A. (1963). *The Method of Paired Comparisons.* New York, NY: Hafner.

Desjardins, J. L., and Doherty, K. A. (2013). Age-related changes in listening effort for various types of masker noises. *Ear Hear.* 34, 261–272. doi: 10.1097/AUD.0b013e31826d0ba4

Dillon, H. (2012). *Hearing Aids.* Sydney, NSW: Boomerang Press.

Elko, G. W., and Pong, A.-T. N. (1995). "A simple adaptive first-order differential microphone," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (New Paltz, NY), 169–172.

Fidell, L. S., and Tabachnick, B. G. (2006). *Using Multivariate Statistics.* New York, NY: Harper and Row.

Gatehouse, S., and Akeroyd, M. A. (2008). The effects of cueing temporal and spatial attention on word recognition in a complex listening task in hearing-impaired listeners. *Trends Amplif.* 12, 145–161. doi: 10.1177/1084713808317395

Gatehouse, S., Naylor, G., and Elberling, C. (2006a). Linear and nonlinear hearing aid fittings - 1. Patterns of benefit. *Int. J. Audiol.* 45, 130–152. doi: 10.1080/14992020500429518

Gatehouse, S., Naylor, G., and Elberling, C. (2006b). Linear and nonlinear hearing aid fittings - 2. Patterns of candidature. *Int. J. Audiol.* 45, 153–171. doi: 10.1080/14992020500429484

Glista, D., Scollie, S., Bagatto, M., Seewald, R., Parsa, V., and Johnson, A. (2009). Evaluation of nonlinear frequency compression: clinical outcomes. *Int. J. Audiol.* 48, 632–644. doi: 10.1080/14992020902971349

Grimm, G., Herzke, T., Berg, D., and Hohmann, V. (2006). The master hearing aid: a PC-based platform for algorithm development and evaluation. *Acta Acust. Acust.* 92, 618–628.

Houben, R., Dijkstra, T. M. H., and Dreschler, W. A. (2012). "The influence of noise type on the preferred setting of a noise reduction algorithm," in *Speech Perception and Auditory Disorders*, eds T. Dau, M. L. Jepsen, T. Poulsen, and J. C. Dalsgaard (Nyborg: The Danavox Jubilee Foundation), 465–472.

Jensen, N. S., Neher, T., Laugesen, S., Johannesson, R. B., and Kragelund, L. (2013). Laboratory and field study of the potential benefits of pinna cue-preserving hearing aids. *Trends Amplif.* 17, 171–188. doi: 10.1177/1084713813510977

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M., et al. (2004). DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *Int. J. Geriatr. Psychiatry* 19, 136–143. doi: 10.1002/gps.1042

Kates, J. M., and Arehart, K. H. (2014). The Hearing-Aid Speech Quality Index (HASQI) Version 2. *J. Audio Eng. Soc.* 62, 99–117.

Kayser, H., Ewert, S. D., Anemüller, J., Rohdenburg, T., Hohmann, V., and Kollmeier, B. (2009). Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP J. Adv. Signal Process.* 2009:298605. doi: 10.1155/2009/298605

Keidser, G., Dillon, H., Convery, E., and Mejia, J. (2013). Factors influencing individual variation in perceptual directional microphone benefit. *J. Am. Acad. Audiol.* 24, 955–968. doi: 10.3766/jaaa.24.10.7

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42(Suppl. 1), S49–S58. doi: 10.3109/14992020309074624

Lunner, T., Rudner, M., and Rönnberg, J. (2009). Cognition and hearing aids. *Scand. J. Psychol.* 50, 395–403. doi: 10.1111/j.1467-9450.2009.00742.x

Lunner, T., and Sundewall-Thorén, E. (2007). Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J. Am. Acad. Audiol.* 18, 604–617. doi: 10.3766/jaaa.18.7.7

Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M., Buechler, M., et al. (2010). Multicenter evaluation of signal enhancement algorithms for hearing aids. *J. Acoust. Soc. Am.* 127, 1491–1505. doi: 10.1121/1.3299168

Neher, T., Grimm, G., and Hohmann, V. (2014). Perceptual consequences of different signal changes due to binaural noise reduction: do hearing loss and working memory capacity play a role? *Ear Hear.* 35, e213–e227. doi: 10.1097/AUD.0000000000000054

Neher, T., Grimm, G., Hohmann, V., and Kollmeier, B. (2013). Do hearing loss and cognitive function modulate benefit from different binaural noise-reduction settings? *Ear Hear.* 53, e52–e62. doi: 10.1097/AUD.0000000000000003

Neher, T., Laugesen, S., Jensen, N. S., and Kragelund, L. (2011). Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *J. Acoust. Soc. Am.* 130, 1542–1558. doi: 10.1121/1.3608122

Neher, T., Lunner, T., Hopkins, K., and Moore, B. C. J. (2012). Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners (L). *J. Acoust. Soc. Am.* 131, 2561–2564. doi: 10.1121/1.3689850

Ng, E. H., Rudner, M., Lunner, T., Pedersen, M. S., and Rönnberg, J. (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *Int. J. Audiol.* 52, 433–441. doi: 10.3109/14992027.2013.776181

Ricketts, T. A., and Mueller, H. G. (2000). Predicting directional hearing aid benefit for individual listeners. *J. Am. Acad. Audiol.* 11, 561–569.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42(Suppl. 1), S68–S76. doi: 10.3109/14992020309074626

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S99–S105. doi: 10.1080/14992020802301167

Smeds, K., Bergman, N., Hertzman, S., and Nyman, T. (2010). "Noise reduction in modern hearing aids—Long-term average gain measurements using speech," in *Binaural Processing and Spatial Hearing*, eds J. M. Buchholz, T. Dau, J. Cristensen-Dalsgaard, and T. Poulsen (Elsinore: The Danavox Jubilee Foundation), 445–452.

Souza, P. E., Arehart, K. H., Kates, J. M., Croghan, N. B., and Gehani, N. (2013). Exploring the limits of frequency lowering. *J. Speech Lang. Hear. Res.* 56, 1349–1363. doi: 10.1044/1092-4388(2013/12-0151)

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J. Speech Hear. Res.* 28, 455–462. doi: 10.1044/jshr.2803.455

Van Breukelen, G. J. P., and Van Dijk, K. R. A. (2007). Use of covariates in randomized controlled trials. *J. Int. Neuropsychol. Soc.* 13, 903–904. doi: 10.1017/S1355617707071147

Van Den Bogaert, T., Wouters, J., Klasen, T. J., and Moonen, M. (2005). "Distortion of interaural time cues by directional noise reduction systems in modern digital hearing aids," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (New Paltz, NY), 57–60. doi: 10.1109/ASPAA.2005.1540167

Wagener, K., Brand, T., and Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache. I-III: design, Optimierung und Evaluation des Oldenburger Satztests (Development and evaluation of a sentence test for the German language. I-III: Design, optimization and evaluation of the Oldenburg sentence test). *Zeitschrift für Audiologie (Audiological Acoustics)* 38, 4–15, 44–56, 86–95.

Walden, B. E., Surr, R. K., Cord, M. T., and Dyrlund, O. (2004). Predicting hearing aid microphone preference in everyday listening. *J. Am. Acad. Audiol.* 15, 365–396. doi: 10.3766/jaaa.15.5.4

Walden, B. E., Surr, R. K., Grant, K. W., Van Summers, W., Cord, M. T., and Dyrlund, O. (2005). Effect of signal-to-noise ratio on directional microphone benefit and preference. *J. Am. Acad. Audiol.* 16, 662–676. doi: 10.3766/jaaa.16.9.4

Zimmermann, P., and Fimm, B. (2012). *Testbatterie zur Aufmerksamkeitsprüfung - Version Mobilität (Test battery for the assessment of attentional skills—Mobility version)*. Herzogenrath: Psytest.

Conflict of Interest Statement: The research reported in this article was co-funded by Siemens Audiologische Technik, Erlangen, Germany. However, the contents represent the work and private views of the author only.

Received: 29 September 2014; accepted: 14 November 2014; published online: 04 December 2014.

Citation: Neher T (2014) Relating hearing loss and executive functions to hearing aid users' preference for, and speech recognition with, different combinations of binaural noise reduction and microphone directionality. Front. Neurosci. ***8****:391. doi: 10.3389/fnins.2014.00391*

This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Neuroscience.

Copyright © 2014 Neher. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

PERSPECTIVE
published: 28 May 2015
doi: 10.3389/fpsyg.2015.00556

Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss

Melanie A. Ferguson[1,2]* *and Helen Henshaw*[1]

[1] NIHR Nottingham Hearing Biomedical Research Unit, Otology and Hearing Group, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, UK, [2] Nottingham University Hospitals NHS Trust, Nottingham, UK

Auditory training (AT) helps compensate for degradation in the auditory signal. A series of three high-quality training studies are discussed, which include, (i) a randomized controlled trial (RCT) of phoneme discrimination in quiet that trained adults with mild hearing loss ($n = 44$), (ii) a repeated measures study that trained phoneme discrimination in noise in hearing aid (HA) users ($n = 30$), and (iii) a double-blind RCT that directly trained working memory (WM) in HA users ($n = 57$). AT resulted in generalized improvements in measures of self-reported hearing, competing speech, and complex cognitive tasks that all index executive functions. This suggests that for AT related benefits, the development of complex cognitive skills may be more important than the refinement of sensory processing. Furthermore, outcome measures should be sensitive to the functional benefits of AT. For WM training, lack of far-transfer to untrained outcomes suggests no generalized benefits to real-world listening abilities. We propose that combined auditory-cognitive training approaches, where cognitive enhancement is embedded within auditory tasks, are most likely to offer generalized benefits to the real-world listening abilities of adults with hearing loss.

Keywords: auditory training, hearing loss, working memory, attention, communication, hearing aids, executive function, speech perception

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Larry E. Humes,
Indiana University Bloomington, USA
Paula Clare Stacey,
Nottingham Trent University, UK

***Correspondence:**
Melanie A. Ferguson,
NIHR Nottingham Hearing Biomedical Research Unit, Otology and Hearing Group, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Ropewalk House, 113 Ropewalk, Nottingham, NG1 5DU, UK
melanie.ferguson@nottingham.ac.uk

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *11 February 2015*
Accepted: *16 April 2015*
Published: *28 May 2015*

Citation:
Ferguson MA and Henshaw H (2015) Auditory training can improve working memory, attention, and communication in adverse conditions for adults with hearing loss. Front. Psychol. 6:556. doi: 10.3389/fpsyg.2015.00556

Listening and Communication in Adverse Conditions

It is widely accepted that understanding speech in background noise is the most common problem for people with hearing loss (Vermiglio et al., 2012; Humes et al., 2013), as characterized by the typical statement "I can hear but I cannot understand what is being said." In addition to a loss of hearing sensitivity, there may be additional deficits of temporal and spectral processing that contribute to listening difficulties (Hopkins and Moore, 2011). Furthermore, there is mounting evidence that non-sensory factors such as cognition, motivation, and context, play an important role in both listening to speech (one-way interaction process) and communication (bi-directional interaction; Kiessling et al., 2003; Pichora-Fuller and Singh, 2006; Rudner et al., 2011). This is particularly evident for older listeners (Gordon-Salant, 2014; Moore et al., 2014).

The role of cognition becomes more apparent when communicating in adverse conditions, such as when listening to speech in fluctuating background noise or competing talkers

(Akeroyd, 2008; Humes and Dubno, 2010). Speech in noise performance is associated with cognition, and the role of cognition becomes increasingly important as the complexity of the listening task increases (Heinrich et al., 2015). For a listener to be able to understand a specific speech source amongst a background of other talkers, the auditory streams or sound sources need to be simultaneously attended to and monitored, and attention may need to be switched between them (Gatehouse and Noble, 2004; Shinn-Cunningham and Best, 2008). This requires the engagement of executive processes that regulate, control, and manage other cognitive processes, such as attention and working memory (WM; Chan et al., 2008).

Cognition and the Clinical Management of People with Hearing Loss

The role of cognition has implications for the clinical management of people with hearing loss. Hearing aids (HAs) are the main intervention for people with hearing loss and have undergone significant advances in digital technology over the last two decades. Whilst satisfaction with HAs has improved since the 1990s (Kochkin, 2010), users often continue to encounter difficulties in challenging listening conditions (Johnson and Dillon, 2011). Early studies with HA users showed an association between behavioral and subjective HA outcomes and measures of cognitive skills (Gatehouse et al., 2003; Lunner, 2003). Furthermore, those with better cognitive skills were better able to take advantage of advanced signal processing strategies, such as fast-acting compression (Foo et al., 2007; Lunner and Sundewall-Thorén, 2007). Other processing strategies, such as noise reduction algorithms, have also been shown to reduce effortful listening and free up cognitive resources to be used for other tasks (Sarampalis et al., 2009).

When considering interventions to aid communication in people with hearing loss, HAs alone are not the only option. Other rehabilitation approaches include patient-centered education, counseling, and auditory perceptual training, which can help impaired listeners compensate for degradation in the auditory signal and improve communication (Sweetow and Sabes, 2006). This article focusses on developments in auditory training (AT), and more recently cognitive training, and how this may improve speech perception, cognition and ultimately, everyday communication in adults with hearing loss, offering a view to future research directions.

Auditory Training

Auditory perceptual training can be described as teaching the brain to listen through active engagement with sounds, whereby listeners typically learn to make perceptual distinctions between sounds presented systematically (Schow and Nerbonne, 2006). Training on perceptual distinctions implies a primarily bottom-up approach to training whereby the individual actively listens to auditory stimuli (e.g., tones, phonemes, words) to improve listening and communication. This is reflected in the literature where traditionally, training studies have focussed primarily on the sensory refinement of auditory stimuli to improve speech perception (Fu et al., 2004; Stecker et al., 2006). But as Schow and Nerbonne's (2006) definition suggests, the role of top–down cognitive processes is implicit in AT and subsequent learning. This has been demonstrated by training on non-auditory tasks, such as visual discrimination or visuospatial tasks, and auditory tasks with identical stimuli, resulting in learning in the auditory domain (Amitay et al., 2006). Such results imply that learning is mediated by top–down processes. Thus, AT may provide a means to improve both auditory and cognitive processes in people with hearing loss in order to improve listening and communication in everyday life (Pichora-Fuller and Levitt, 2012).

Efficacy of Auditory Training

The turn of the last decade saw a proliferation of individualized, computer-based auditory training research. Basic research sought to better understand the underlying principles and biological mechanisms of AT in normally hearing listeners (e.g., Tremblay, 2007; de Boer and Thornton, 2008; Wright and Zhang, 2009; Song et al., 2011). In addition, translational research sought to establish the efficacy of AT to improve outcomes for people with hearing loss, including users of HAs and cochlear implants (for review, see Henshaw and Ferguson, 2013a). Efficacy of AT can be assessed by (i) improvements in performance for the trained task (on-task learning), (ii) improvements in performance on the untrained task (off-task, generalized, or transfer of learning), (iii) retention of learning for a period after training ceases, and (iv) adherence of the individual with training. This article concentrates on (i)–(iii). Motivations of individuals to participate in, engage with, and adhere to home-delivered training are discussed elsewhere (Henshaw et al., in review; Ferguson and Henshaw, in press).

Our recent systematic review on the efficacy of computer-based auditory training as a clinical intervention for adults with hearing loss summarized the evidence base between 1996 and 2011 and included 13 studies (Henshaw and Ferguson, 2013a). The review concluded that, where reported, on-task learning always occurred in those with mild-moderate hearing loss (whether HA users or not) for a range of training stimuli including phonemes, words, and sentences (e.g., Burk et al., 2006; Stecker et al., 2006; Sweetow and Sabes, 2006). The evidence for on-task learning in cochlear implants users generally followed this trend (e.g., Fu et al., 2004; Tyler et al., 2010; Oba et al., 2011) with the exception of Stacey et al. (2010). However, the evidence for generalization of learning to untrained measures was mixed. Although generalized improvements were shown for speech intelligibility (11/13 articles), self-report of communication (1/2), and cognition (1/1), the improvements were variable in that reported improvements were inconsistent across studies, and the magnitude of improvement was small and not robust. It was notable that all the studies had at least one outcome measure on speech intelligibility, yet different studies rarely used the same measure. Only two studies measured self-reported communication as a means to tap into perceived real-world benefits of training, and just one study measured cognition. The quality of the evidence for included studies was very-low to moderate. Reasons for this included failure to include a control

group, and a lack of randomization, power calculation, and participant or tester blinding.

Our Approach to Auditory Training

Following on from the systematic review, we sought to address many of the study quality limitations of the existing published evidence with a series of three high-quality auditory and cognitive training studies that aimed to assess benefits to speech perception, cognition, and self-reported communication in people with mild-moderate hearing loss. The study methods are outlined in **Table 1**. Outcome measures are shown in **Table 2**, and are described in more detail in the original articles.

Across all three studies, hearing loss was described by the better-ear pure-tone threshold averaged across octave frequencies 0.5–4 kHz as either mild (21–40 dB HL), or moderate (41–70 dB HL). Participants were aged 50–74 years old, and training was home-delivered either via loan laptops (AT studies) or via the internet (working memory training). Each study included a control period that allowed for the examination of procedural learning (test–retest) effects on outcomes (Mcarthur, 2007).

Auditory Training Study 1: Training Improves Outcomes that Index Executive Function

The study was a randomized controlled trial, whereby a 4-week phoneme discrimination training program was performed for the Immediate Trained (IT) group at weeks 1–4, and a Delayed Trained (DT) group at weeks 5–8 provided a control comparison. Outcome measure assessments were obtained for the IT and DT groups at weeks 0, 4, and 8, and for the DT group at 12 weeks (Ferguson et al., 2014).

Results showed significant and robust on-task learning for all trained phoneme continua. The on-task learning and retention of on-task learning results were consistent with studies in the systematic review. However, from a clinical perspective the value of training as an intervention lies in the generalization of task-specific learning to functional benefits in real-world listening. A summary of the results from the untrained outcome measures is shown in **Table 2**, whereby tests and self-report questions were classified as complex if they indexed executive processes, and simple if they did not. Details of analysis using Multivariate Analysis of Variance is reported elsewhere (Ferguson et al., 2014). As we were also interested in the clinical effects of AT as an intervention, Cohen's *d* is reported where effect size was interpreted as small (0.2), moderate (0.5), and large (0.8) (Cohen, 1988).

For the speech perception in noise tests that used energetic masking, there were no significant training-related improvements. For tests of cognition, there were no pre–post training improvements for the simple tasks, including simple-span WM measure (digit span) and the single attention test [Test of Everyday Attention (TEA) subtest 6] for either the intervention or control groups. However, for the complex tasks that indexed executive processes, there were significant pre–post training improvements shown for divided attention (TEA subtest 7) and

TABLE 1 | Study and participant characteristics.

Methods	Auditory training 1 Ferguson et al. (2014)	Auditory training 2 Henshaw and Ferguson (2014)	Working memory training Henshaw and Ferguson (2013b)
Study design	Randomized controlled trial	Repeated measures	Randomized controlled trial
Intervention	Phoneme discrimination in quiet 11 phoneme continua Phonomena/IHR STAR	Phoneme discrimination in multitalker babble 11 phoneme continua Phonomena/IHR Star	Verbal and Visuospatial working memory and storage tasks Cogmed RM
Intervention duration requested	360 min (6 h) across 4 weeks 15 min/day, 6 days/week (total = 24 sessions) Immediate Trained (IT): weeks 1–4 Delayed Training (DT): weeks 5–8	210 min (3.5 h) across 1 week (7 days), 2 × 15 min/day (total = 14 sessions)	Approximately 990 min (16.5 h) across 5 weeks, 35–45 min sessions/day, 5 days/week (total = 25 sessions)
Retention period	4 weeks post-training: IT: weeks 4–8 DT: weeks 8–12	None	6 months post-training
Control activity	None	None	Active control working memory tasks – span fixed at 3
Control period: duration	DT: T1–T2 = 4 weeks	T1–T2 = 1 week	T2–T3 = 5 weeks
Test–retest period	DT: T1–T2 = 4 weeks	T1–T2 = 1 week	T1–T2 = 1 week
n participants (n females)	44 (15)	30 (10)	57 (30)
Participants (source of recruitment)	Non-HA users, mild hearing loss (general practitioner)	Existing HA users, mild-moderate hearing loss (volunteer database)	Existing HA users, mild-moderate hearing loss (volunteer database)
Mean age in years (SD)	65.3 (5.7)	67.4 (7.1)	64.9 (6.0)
Mean $BEA_{0.5-4\ kHz}$ (SD) dB HL	32.5 (6.0)	43.6 (13.6)	44.0 (13.8)

IT, immediate trained; DT, delayed trained; HA, hearing aid; BEA, better-ear average; min, minute; T1–T2, time period between the first two test sessions to measure test–retest effects.

TABLE 2 | Summary of results for untrained tasks.

	Simple		Complex	
	Test	**Results**	**Test**	**Results**
Auditory training 1: Ferguson et al. (2014) ($n = 44$)				
Speech perception	Digit Triplet Test in steady speech-shaped noise[1]	NS within group for IT or DT ($p > 0.05$)	N/T	N/A
	Adaptive Sentence List in 8-Hz modulated noise[2]	NS within group effect for IT or DT ($p > 0.05$)	N/T	N/A
Cognition	Digit Span[3]	NS within group for IT or DT ($p > 0.05$)	Visual letter monitoring[4] 1/s (fast) updating	All trained: $p = 0.003$, $d = 0.50$ DT control: NS ($p > 0.05$)
	Test of Everyday Attention[5] – single attention task	NS within group for IT or DT ($p > 0.05$)	Test of Everyday Attention – dual task decrement	All trained: $p = 0.001$, $d = 0.53$ DT control: NS ($p > 0.05$)
Communication	Glasgow Hearing Aid Benefit Profile[6]: Television set to set to level for others 1:1 conversation in no background noise Conversation in a busy street or shop	NS within group for IT or DT ($p > 0.05$)	Glasgow Hearing Aid Benefit Profile: Having a conversation with several people in a group	All trained: $p = 0.005$, $d = 0.68$ DT control: NS ($p > 0.05$)
Auditory training 2: Henshaw and Ferguson (2014) ($n = 30$)				
Speech perception	N/A	N/A	Competing speech[7]	Trained period: $p = 0.03$, $d = 0.47$ Control period: NS ($p > 0.05$)
			Dual task of speech and memory[8]	Trained period: $p = 0.001$, $d = 0.77$ Control period: NS ($p > 0.05$)

IT, immediate trained; DT, delayed Trained; NS, no significant effect; N/T, not tested; N/A, not applicable. [1]*Smits et al. (2004),* [2]*Millward et al. (2011),*[3]*Wechsler (1997),* [4]*Gatehouse et al. (2003),* [5]*Robertson et al. (1996),* [6]*Gatehouse (1999),* [7]*Hazan et al. (2009),* [8]*Howard et al. (2010).*

the updating of WM (visual letter monitoring, VLM). For VLM there was a larger effect shown for the faster, more challenging presentation (one letter per second, $d = 0.50$) compared to the slower presentation (one letter every 2 s, $d = 0.34$).

For self-report of communication using the Glasgow Hearing Aid Benefit Profile (GHABP), there was a significant effect of training on the overall score for activity limitation (previously termed hearing disability) with a moderate effect size, suggesting real-world benefits were perceived by participants. A secondary analysis of the individual situations of the GHABP revealed an interesting insight in that no significant pre–post training improvements for the simple listening situations, such as 'having a conversation with one other person when there is no background noise' were shown. However, there was a significant effect of training for the most challenging listening situation 'having a conversation with several people in a group.' This requires the listener to constantly monitor the conversation, switch, and update attention (i.e., engage executive processes), whilst the other situations do not. These results were supported by qualitative analysis of open-ended questions and focus groups from participants who reported that the main benefits of the training were increased concentration and focus in everyday listening (Henshaw et al., in review). Across all measures where there were significant effects of training, these were retained 4 weeks post-training. Finally, in the participants where there were improvements in the GHABP measures, there was a significant correlation between self-report and divided attention ($r = 0.79$, $p < 0.001$), suggesting that improvements in self-report were not a 'placebo' effect of undertaking the training program.

These results suggest that outcome measures need to be appropriately complex and challenging to be sensitive to the effects of AT, and taken together, the value of AT to mediate cognitive skills may be more important than the improvement of sensory skills for communication in everyday life.

This led us to reconsider the non-significant speech perception results. Given that AT showed an improvement in the cognitive functions that index executive processes, we made the hypothesis that training-related improvements would be evident in informational masked speech perception tests (e.g., competing speech) that engage executive processes (Shinn-Cunningham, 2008), rather than the energetically masked speech tests that were included in this study, which primarily assess audibility. This was explored in study 2.

Auditory Training Study 2: Training Improves Competing Speech and Dual-Task Performance

This study used a within-participant repeated measures design with an initial 1-week control period, followed by a 1-week training period (Henshaw and Ferguson, 2014). The training duration was 3.5 h, just over half that of the previous study, as the majority

of the phoneme discrimination learning had taken place by this time. The modified co-ordinate response measure (MCRM) used a single female talker target and single male talker masker, presented simultaneously. The dual-task included a digit recall task (secondary), which flanked a word-in-noise repetition test (primary), presented at three signal-to-noise (SNR) levels (quiet, 0 dB and −4 dB).

Participants demonstrated significant on-task learning for the trained auditory task. Results for the untrained measures are shown in **Table 2**. For competing speech (MCRM), there was a significant pre–post training improvement with a moderate effect size and no improvement shown for the control (no-training) period. This confirmed our hypothesis and suggests that it is important to use appropriate speech measures that tap into the underpinning mechanisms of benefit provided by AT.

For the dual task, there was no effect of training for the easiest (quiet) or most difficult (−4 dB SNR) test conditions. However, there was a significant pre–post training improvement for the intermediate level of difficulty (0 dB SNR), with a large effect size. This suggests that the HA users in this study were better able to allocate their available cognitive resources between the speech and memory tasks post-training, and suggests that outcome measures need to be appropriately challenging in order to be sensitive to post-training improvements.

Given these results, we asked the question: "Could training cognition directly offer a more direct route to benefit for people with hearing loss?."

Working Memory Training for People With Hearing Loss

We used a WM training program (Cogmed RM) comprising verbal and visuospatial WM and memory storage tasks. Published studies of Cogmed RM have shown post-training improvements in untrained tasks of attention and self-report of cognitive function in younger and older adults (Brehmer et al., 2012), and improvements in sentence repetition for children with cochlear implants (Kronenberger et al., 2011).

Working Memory Training: Training Results in Near-Transfer but not Far-Transfer of Learning

A registered clinical trial of 57 existing HA users with mild-moderate hearing loss assessed benefits to speech perception, self-reported communication, and cognition (for protocol, including outcome measures, see Henshaw and Ferguson, 2013b). In addition to assessing generalization to untrained tasks, we examined how far along the spectrum of near-transfer (e.g., outcome is close to the trained task) to far-transfer (e.g., untrained task in a different modality) any improvements occurred (Perkins and Salomon, 1992). Results (not yet published), showed near-transfer (i.e., improvements in an untrained WM task), but no far-transfer (e.g., speech perception) of training-related improvements, despite a longer training duration than for the AT studies. These results are consistent with the cognitive neuroscience literature, which shows that WM training can enhance WM tasks that share similar structural features (Thompson et al., 2013), however, training does not generalize to enhancement of the broader underlying cognitive constructs (Melby-Lervag and Hulme, 2013). It has been suggested that training-related improvements in trained WM tasks may be mediated by specific strategies, such as chunking or grouping (Dunning and Holmes, 2014), which may limit the broader applicability to benefit cognitive constructs underpinning successful communication for HA users.

Auditory-Cognitive Training: Joined-up Listening and Thinking

Recent studies of an auditory-based cognitive training program that combines auditory perceptual training with increased memory demands (Brain Fitness; Posit Science) have demonstrated generalized improvements in non-trained tests of memory, attention, and speed of processing in older adults (Smith et al., 2009), in addition to improved neural timing and speech perception in noise (Anderson et al., 2013a,b). Similar results for a 'hybrid' training program comprising exercises of speech and cognition [Listening and Communication Enhancement (LACE), Sweetow and Sabes, 2006] trialed in mainly HA users, showed generalized improvements in speech in noise, auditory WM and speed of processing, in addition to improvements in self-report of communication difficulties. However, it is not clear from these studies which element of the training program is responsible for the transfer of learning.

Future Directions

Following on from our own research and developments from the current literature, we propose that benefits of training for people with hearing loss in terms of improved speech understanding in adverse conditions may be best achieved if an integrated auditory-cognitive training approach is taken. This approach would serve to target the cognitive processes that underpin speech perception within a speech task, rather than training specific cognitive tasks that are far-removed from speech. One benefit of this approach is that the degree of transfer required to realize real-world benefit is substantially reduced. Furthermore, the nature of the speech task is more readily perceived as relevant to individuals in terms of their hearing difficulties, which is likely to aid motivation for adherence (Henshaw et al., in review).

Finally, a recent study has shown a dynamic relationship between WM capacity and speech recognition in the first 6 months of HA use with WM playing a greater role in speech perception initially, whereas after 6 months hearing sensitivity is more influential (Ng et al., 2014). Based on the Ease of Language Understanding model (Ronnberg et al., 2013), the authors suggest that as the unfamiliar processed phonological representations become more familiar with time, often referred to as acclimatization (Arlinger et al., 1996), there is a reduced requirement to use WM capacity for speech perception. However, the role of

cognition in the acclimatization process is likely to extend beyond WM, and may call upon executive processes required for understanding speech. We are currently examining this in a longitudinal study of first-time HA users. Having identified which cognitive processes are important in acclimatization we aim to use a relevant auditory-cognitive training paradigm to minimize the difficulties faced in the early days of HA use.

Author Contributions

MF and HH designed the studies. MF and HH analyzed and interpreted the data. MF wrote the manuscript. MF and HH contributed to manuscript revisions and critical discussions. Both authors approved the final version of the manuscript for publication. Both authors agree to be accountable for all aspects of the work and in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgments

The authors would like to thank our colleagues Dave Moore, Daniel Clark, Holly Thomas, Ashana Tittle, and Mark Edmondson-Jones for their contributions to this research. Cogmed is a registered trademark of Pearson, Inc. or its affiliate(s). All rights reserved. This paper presents independent research funded by the National Institute for Health Research (NIHR) Biomedical Research Unit Programme. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Amitay, S., Irwin, A., and Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nat. Neurosci.* 9, 1446–1448. doi: 10.1038/nn1787

Anderson, S., White-Schwoch, T., Choi, H. J., and Kraus, N. (2013a). Training changes processing of speech cues in older adults with hearing loss. *Front. Syst. Neurosci.* 7:97. doi: 10.3389/fnsys.2013.00097

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013b). Reversal of age-related neural timing delays with training. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4357–4362.

Arlinger, S., Gatehouse, S., Bentler, R., Byrne, D., Cox, R., Dirks, D., et al. (1996). Report of the eriksholm workshop on auditory deprivation and acclimatization. *Ear Hear.* 17(Suppl. 3), 87S–98S. doi: 10.1097/00003446-199617031-00009

Brehmer, Y., Westerberg, H., and Backman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Front. Hum. Neurosci.* 6:63. doi: 10.3389/fnhum.2012.00063

Burk, M. H., Humes, L. E., Amos, N. E., and Strauser, L. E. (2006). Effect of training on word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners. *Ear Hear.* 27, 263–278. doi: 10.1097/01.aud.0000215980.21158.a2

Chan, R. C. K., Shum, D., Toulopoulou, T., and Chen, E. Y. H. (2008). Assessment of executive functions: review of instruments and identification of critical issues. *Arch. Clin. Neuropsychol.* 23, 201–216. doi: 10.1016/j.acn.2007.08.010

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Earlbaum Associates.

de Boer, J., and Thornton, A. R. D. (2008). Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *J. Neurosci.* 28, 4929–4937. doi: 10.1523/JNEUROSCI.0902-08.2008

Dunning, D. L., and Holmes, J. (2014). Does working memory training promote the use of strategies on untrained working memory tasks? *Mem. Cognit.* 42, 854–862.

Ferguson, M. A., and Henshaw, H. (in press). Computer and internet interventions to optimise listening and learning for people with hearing loss: accessibility, use and adherence. *Am. J. Audiol.*

Ferguson, M. A., Henshaw, H., Clark, D., and Moore, D. (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50–74 year olds with mild hearing loss. *Ear Hear.* 35, e110–e121. doi: 10.1097/AUD.0000000000000020

Foo, C., Rudner, M., Ronnberg, J., and Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *J. Am. Acad. Audiol.* 18, 618–631. doi: 10.3766/jaaa.18.7.8

Fu, Q. J., Galvin, J., Wang, X., and Nogaki, G. (2004). Effects of auditory training on adult cochlear implant patients: a preliminary report. *Cochlear Implants Int.* 5, 84–90. doi: 10.1179/cim.2004.5.Supplement-1.84

Gatehouse, S. (1999). Glasgow hearing aid benefit profile: derivation and validation of client-centred outcome measures for hearing aid services. *J. Am. Acad. Audiol.* 10, 80–103.

Gatehouse, S., Naylor, G., and Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *Int. J. Audiol.* 42, S77–S85. doi: 10.3109/14992020309074627

Gatehouse, S., and Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *Int. J. Audiol.* 43, 85–99. doi: 10.1080/14992020400050014

Gordon-Salant, S. (2014). "Aging, hearing loss, and speech recognition: stop shouting, i can't understand you," in *Perspectives on Auditory Research*, eds A. N. Popper and R. R. Fay (New York, NY: Springer), 211–228.

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., and Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: is there any evidence for a true deficit? *J. Speech Lang. Hear. Res.* 52, 1510–1529. doi: 10.1044/1092-4388(2009/08-0220)

Heinrich, A., Henshaw, H., and Ferguson, M. A. (2015). The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front. Psychol.* 6:782. doi: 10.3389/fpsyg.2015.00782

Henshaw, H., and Ferguson, M. A. (2013a). Efficacy of individual computer-based auditory training for people with hearing loss: a systematic review of the evidence. *PLoS ONE* 8:e62836. doi: 10.1371/journal.pone.0062836

Henshaw, H., and Ferguson, M. A. (2013b). Working memory training for adult hearing aid users: study protocol for a double-blind randomized active controlled trial. *Trials* 14:417.

Henshaw, H., and Ferguson, M. A. (2014). "Assessing the benefits of auditory training to real-world listening: identifying appropriate and sensitive outcomes," in *Proceedings of ISAAR 2013: Auditory Plasticity – Listening with the Brain. 4th symposium on Auditory and Audiological Research*, eds T. Dau, S. Santurette, J. C. Dalsgaard, L. Trangjaerg, T. Andersen, and T. Poulsen (Nyborg: The Danavox Jubilee Foundation).

Hopkins, K., and Moore, B. C. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J. Acoust. Soc. Am.* 130, 334–349. doi: 10.1121/1.3585848

Howard, C. S., Munro, K. J., and Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *Int. J. Audiol.* 49, 928–932. doi: 10.3109/14992027.2010.520036

Humes, L. E., and Dubno, J. R. (2010). "Factors affecting speech understanding in older adults," in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisna, A. N. Popper, and R. R. Fay (New York, NY: Springer), 211–257.

Humes, L. E., Kidd, G. R., and Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Front. Syst. Neurosci.* 7:55. doi: 10.3389/fnsys.2013.00055

Johnson, E. E., and Dillon, H. (2011). A comparison of gain for adults from generic hearing aid prescriptive methods: impacts on predicted loudness, frequency bandwidth, and speech intelligibility. *J. Am. Acad. Audiol.* 22, 441–459. doi: 10.3766/jaaa.22.7.5

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., et al. (2003). Candidature for and delivery of audiological services: special needs of older people. *Int. J. Audiol.* 42, S92–S101. doi: 10.3109/14992020309074650

Kochkin, S. (2010). MarkeTrak VIII: consumer satisfaction with hearing aids is slowly increasing. *Hear. J.* 63, 19–20. doi: 10.1097/01.HJ.0000366912.40173.76

Kronenberger, W. G., Pisoni, D. B., Henning, S. C., Colson, B. G., and Hazzard, L. M. (2011). Working memory training for children with cochlear implants: a pilot study. *J. Speech Lang. Hear. Res.* 54, 1182–1196. doi: 10.1044/1092-4388(2010/10-0119)

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *Int. J. Audiol.* 42, S49–S58. doi: 10.3109/14992020309074624

Lunner, T., and Sundewall-Thorén, E. (2007). Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J. Am. Acad. Audiol.* 18, 604–617. doi: 10.3766/jaaa.18.7.7

Mcarthur, G. (2007). Test–retest effects in treatment studies of reading disability: the devil is in the detail. *Dyslexia* 13, 240–252. doi: 10.1002/dys.355

Melby-Lervag, M., and Hulme, C. (2013). Is working memory training effctive? A meta- analytic review. *Dev. Psychol.* 49, 270–291. doi: 10.1037/a0028228

Millward, K., Moore, D. R., Sohoglu, E., and Amitay, S. (2011). Training speech-in-noise perception in mainstream school children. *Int. J. Pediatr. Otorhinolaryngol.* 75, 1408–1417. doi: 10.1016/j.ijporl.2011.08.003

Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., Mccormack, A., Pierzycki, R. H., et al. (2014). Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. *PLoS ONE* 9:e107720. doi: 10.1371/journal.pone.0107720

Ng, E. H., Classon, E., Larsby, B., Arlinger, S., Lunner, T., Rudner, M., et al. (2014). Dynamic relation between working memory capacity and speech recognition in noise during the first 6 months of hearing aid use. *Trends Hear.* 18, 2331216514558688. doi: 10.1177/2331216514558688

Oba, S. I., Fu, Q., and Galvin, J. J. (2011). Digit training in noise can improve cochlear implant users' speech understanding in noise. *Ear Hear.* 32, 573–581. doi: 10.1097/AUD.0b013e31820fc821

Perkins, D. N., and Salomon, G. (1992). "Transfer of learning," in *International Encyclopedia of Education*, 2nd Edn, eds T. Husén and T. Postlethwaite (Oxford: Pergamon Press).

Pichora-Fuller, M. K., and Levitt, H. (2012). Speech comprehension training and auditory and cognitive processing in older adults. *Am. J. Audiol.* 21, 351–357. doi: 10.1044/1059-0889(2012/12-0025)

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Robertson, I. H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *J. Int. Neuropsychol. Soc.* 2, 525–534. doi: 10.1017/S1355617700001697

Ronnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rudner, M., Rönnberg, J., and Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *J. Am. Acad. Audiol.* 22, 156–167. doi: 10.3766/jaaa.22.3.4

Sarampalis, A., Kalluri, S., Edwards, B., and Hafter, E. (2009). Objective measures of listening effort: effects of background noise and noise reduction. *J. Speech Lang. Hear. Res.* 52, 1230–1240. doi: 10.1044/1092-4388(2009/08-0111)

Schow, R., and Nerbonne, M. (2006). *Introduction to Audiologic Rehabilitation*. Boston, MA: Pearson Education.

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends Cogn. Sci.* 12, 182–186. doi: 10.1016/j.tics.2008.02.003

Shinn-Cunningham, B. G., and Best, V. (2008). Selective attention in normal and impaired hearing. *Trends Amplif.* 12, 283–299. doi: 10.1177/1084713808325306

Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., et al. (2009). A cognitive training program based on principles of brain plasticity: results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *J. Am. Geriatr. Soc.* 57, 594–603. doi: 10.1111/j.1532-5415.2008.02167.x

Smits, C., Kapteyn, T. S., and Houtgast, T. (2004). Development and validation of an automatic speech-in-noise screening test by telephone. *Int. J. Audiol.* 43, 15–28. doi: 10.1080/14992020400050004

Song, J. H., Skoe, E., Banai, K., and Kraus, N. (2011). Training to improve hearing speech in noise: biological mechanisms. *Cereb. Cortex* 22, 1180–1190. doi: 10.1093/cercor/bhr196

Stacey, P. C., Raine, C. H., O'donoghue, G. M., Tapper, L., Twomey, T., and Summerfield, A. Q. (2010). Effectiveness of computer-based auditory training for adult users of cochlear implants. *Int. J. Audiol.* 49, 347–356. doi: 10.3109/14992020903397838

Stecker, G. C., Bowman, G. A., Yund, E. W., Herron, T. J., Roup, C. M., and Woods, D. L. (2006). Perceptual training improves syllable identification in new and experienced hearing aid users. *J. Rehabil. Res. Dev.* 43, 537–551. doi: 10.1682/JRRD.2005.11.0171

Sweetow, R. W., and Sabes, J. H. (2006). The need for and development of an Adaptive Listening and Communication Enhancement (LACE) Program. *J. Am. Acad. Audiol.* 17, 538–558. doi: 10.3766/jaaa.17.8.2

Thompson, T. W., Waskom, M. L., Garel, K. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., et al. (2013). Failure of working memory training to enhance cognition or intelligence. *PLoS ONE* 8:e63614. doi: 10.1371/journal.pone.0063614

Tremblay, K. (2007). Training-related changes in the brain: evidence from human auditory-evoked potentials. *Semin. Hear.* 28, 120–132. doi: 10.1055/s-2007-973438

Tyler, R. S., Witt, S. A., Dunn, C. C., and Wang, W. (2010). Initial development of a spatially separated speech-in-noise and localization training program. *J. Am. Acad. Audiol.* 21, 390–403. doi: 10.3766/jaaa.21.6.4

Vermiglio, A. J., Soli, S. D., Freed, D. J., and Fisher, L. M. (2012). The relationship between high-frequency pure-tone hearing loss, hearing in noise test (HINT) thresholds, and the articulation index. *J. Am. Acad. Audiol.* 23, 779–788. doi: 10.3766/jaaa.23.10.4

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale*, 3rd Edn. San Antonio, TX: The Psychological Corporation.

Wright, B. A., and Zhang, Y. (2009). A review of the generalization of auditory learning. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 301–311. doi: 10.1098/rstb.2008.0262

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Ferguson and Henshaw. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 August 2015
doi: 10.3389/fpsyg.2015.01067

Intrinsic and extrinsic motivation is associated with computer-based auditory training uptake, engagement, and adherence for people with hearing loss

Helen Henshaw[1]*, *Abby McCormack*[1] and *Melanie A. Ferguson*[1,2]

[1] Otology and Hearing Group, National Institute for Health Research Nottingham Hearing Biomedical Research Unit, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, UK, [2] Nottingham University Hospitals NHS Trust, Nottingham, UK

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Anna Stigsdotter Neely,
Umeå University, Sweden
Håkan Hua,
Linköping University, Sweden

***Correspondence:**
Helen Henshaw,
NIHR Nottingham Hearing Biomedical Research Unit, Ropewalk House, 113 The Ropewalk, Nottingham, NG1 5DU, UK
helen.henshaw@nottingham.ac.uk

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *11 February 2015*
Accepted: *13 July 2015*
Published: *06 August 2015*

Citation:
Henshaw H, McCormack A and Ferguson MA (2015) Intrinsic and extrinsic motivation is associated with computer-based auditory training uptake, engagement, and adherence for people with hearing loss. Front. Psychol. 6:1067. doi: 10.3389/fpsyg.2015.01067

Hearing aid intervention typically occurs after significant delay, or not at all, resulting in an unmet need for many people with hearing loss. Computer-based auditory training (CBAT) may provide generalized benefits to real-world listening, particularly in adverse listening conditions, and can be conveniently delivered in the home environment. Yet as with any intervention, adherence to CBAT is critical to its success. The main aim of this investigation was to explore motivations for uptake, engagement and adherence with home-delivered CBAT in a randomized controlled trial of adults with mild sensorineural hearing loss (SNHL), with a view to informing future CBAT development. A secondary aim examined perceived benefits of CBAT. Participants ($n = 44$, 50–74 years olds with mild SNHL who did not have hearing aids) completed a 4-week program of phoneme discrimination CBAT at home. Participants' experiences of CBAT were captured using a post-training questionnaire ($n = 44$) and two focus groups ($n = 5$ per group). A mixed-methods approach examined participants' experiences with the intervention, the usability and desirability of the CBAT software, and participants' motivations for CBAT uptake, engagement and adherence. Self-Determination Theory (SDT) was used as a theoretical framework for the interpretation of results. Participants found the CBAT intervention easy to use, interesting and enjoyable. Initial participation in the study was associated with extrinsic motivation (e.g., hearing difficulties). Engagement and adherence with CBAT was influenced by intrinsic (e.g., a desire to achieve higher scores), and extrinsic (e.g., to help others with hearing loss) motivations. Perceived post-training benefits included better concentration and attention leading to improved listening. CBAT also prompted further help-seeking behaviors for some individuals. We see this as an important first-step for informing future theory-driven development of effective CBAT interventions.

Keywords: auditory training, motivation, engagement, adherence, hearing loss, sensorineural, Self-Determination Theory

Introduction

In 2008, the World Health Organization estimated that over 360 million people worldwide had a disabling hearing loss. These figures are expected to rise substantially in the future due to aging of the global population (World Health Organization, 2008). Hearing loss currently affects more than 10 million adults in the UK alone (Action on Hearing Loss, 2011), which corresponds to approximately one in six of the population. The most common management strategy for hearing loss is the provision of hearing aids, which primarily help restore audibility. However, just one in three people who could benefit from hearing aids in the UK actually have them (Davis et al., 2007), resulting in an unmet need for an estimated four million people. For those who do seek intervention, this process takes an average of 10 years, which may in part be related to the fact that 47% of those individuals who report hearing difficulties to their family doctor fail to receive an onward referral to Audiology services (Davis et al., 2007). A scoping review of the literature suggests that of those individuals who do seek audiological intervention and are fitted with hearing aids, between 4 and 24% choose not to use them (McCormack and Fortnum, 2013). In addition, many choose not to use them regularly (Whitmer et al., 2014). Untreated hearing loss can lead to numerous social and emotional issues, including; difficulties with work, social withdrawal, isolation, and depression (Davis et al., 2007). In addition, recent findings from a large cohort study identified an association between hearing loss and incident dementia, whereby the risk increases with the degree of impairment (Lin et al., 2011a,b). Auditory (re)habilitation is however, much wider than the provision of hearing aids alone (Ferguson and Henshaw, 2015b).

One of the most common complaints of people with hearing loss is difficulty listening to speech in the presence of distractors, such as competing talkers or background noise (e.g., Pichora-Fuller and Singh, 2006). In recent years, the role of top-down (cognitive) processes in listening have been subject to rigorous examination that is sufficient to warrant its own field of research, Cognitive Hearing Science (Arlinger et al., 2009; Rönnberg et al., 2011). Speech in noise performance is associated with cognition, and the role of cognition becomes increasingly important as the complexity of the listening task increases (Heinrich et al., 2015). Auditory training is one type of intervention for those with hearing loss, which can be described as teaching the brain to listen through active engagement with sounds (Schow and Nerbonne, 2006). Auditory training is designed to improve an individuals' use of their residual hearing through repeated listening practice (Tye-Murray et al., 2012). Both basic and applied research has identified top-down influences of auditory training (Amitay et al., 2006; Sweetow and Henderson Sabes, 2006; Pichora-Fuller and Levitt, 2012; Anderson et al., 2013). For example, Amitay et al. (2006) show that participants are better able to discriminate tone frequency after training on a task that uses identical frequency stimuli. The authors attribute this post-training improvement to both bottom-up and top-down influences, including selective attention and arousal. Evidence from our own research takes this further by suggesting that the benefits of auditory training may be primarily driven by top–down mechanisms, and that these benefits are most evident for challenging listening conditions that index executive processes such as the updating of working memory and attentional control (Ferguson and Henshaw, 2015a,b). These conclusions are based on the results of a randomized controlled trial of 44 adults with mild bilateral sensorineural hearing loss (SNHL) (Ferguson et al., 2014). Post-training outcomes showed significant improvements in divided attention, updating of working memory, and self-reported hearing abilities in a challenging listening condition ("talking with several people in a group") for the trained group, with no improvements for the control group, following a 4-week at-home phoneme discrimination training program (total training time = 6 h). There were no significant improvements shown for a sentence in noise perception task. However, a second study assessing cognitively demanding listening tasks showed a significant improvement for a two-competing talker task (Modified Coordinate Response Measure) of 2.3 dB signal to noise ratio (SNR), following just 3.5 h phoneme discrimination in noise training (Henshaw and Ferguson, 2014).

Yet, as with any intervention, auditory training can only ever be effective if adhered to. In a recent systematic review of 13 articles assessing the efficacy of individual computer-based auditory training (CBAT) for people with hearing loss (Henshaw and Ferguson, 2013), compliance with CBAT was reported in less than half (6/13) of the studies. Where it was reported, compliance rates were high for both laboratory-based (81%) and home-based (73–100%) interventions. However, variation in the definition of training compliance was a particular issue highlighted by the systematic review, with authors often choosing to report either the proportion of participants who did not drop out of the study (i.e., study completion), or the proportion of participants who completed the recommended amount of training (intervention compliance). Sweetow and Henderson Sabes (2006) argued that for widespread use in adults with hearing loss, CBAT programs should be easy, fun, and rewarding, incorporating both top-down and bottom-up approaches to auditory learning. However, data collected from over 3000 patients in routine clinical practice who used the Listening and Communication Enhancement (LACE) CBAT program showed compliance rates of >30%, where compliance was defined as completion of 10 or more of the 20 recommended training sessions (Sweetow, 2009). In another study using LACE, 50 veterans with hearing loss who completed the 20 recommended sessions gained generalized benefits in untrained measures of rapid speech and speech understanding in noise, with no significant improvements for non-compliers (Chisolm et al., 2013). Some of the key challenges associated with auditory training adherence are thought to include; lack of recommendation by audiologists, the nature of the trained task, and the misalignment of audiologist and patient goals (Sweetow and Henderson Sabes, 2010). Nevertheless, these challenges have yet to be confirmed with evidence. In 2003, The World Health Organization placed strong emphasis on the need to differentiate the terms compliance and adherence (World Health Organization, 2003). The main difference is that adherence requires the patient's agreement to the recommendations set, whereas compliance may be more closely associated with blame.

Human behavior is the largest source of variance in health-related outcomes (Schroeder, 2007). Literature from chronic health domains suggests that individuals' motivations play a significant role in treatment adherence (Vermeire et al., 2001). Motivation controls and sustains goal-directed behaviors, with three main components; activation (the decision to initiate the behavior), persistence (continued effort toward a goal even though obstacles may exist), and intensity (the concentration and vigor that goes into pursuing a goal). For home-based training interventions that extend over a number of weeks, there are likely to be several motivational factors that impact on individuals' levels of engagement and adherence with the intervention (Sweetow and Henderson Sabes, 2010). Behavioral science offers opportunities to develop and advance digital health interventions (Pagoto and Bennett, 2013), whereby insights from health behavior psychology can improve our understanding of auditory training adherence and highlight consideration for future auditory training development (Tye-Murray et al., 2012).

Self-Determination Theory (SDT; Deci and Ryan, 1985) is an approach to motivation that is concerned with supporting people's natural tendencies to behave in effective and healthy ways. SDT distinguishes between different types of motivation based on the different reasons or goals that give rise to an action. The basic distinction is between intrinsic motivation, which refers to doing something because it is inherently interesting or enjoyable, and extrinsic motivation, which refers to doing something because it leads to a separable outcome (Ryan and Deci, 2000). As such, intrinsic motivation is important for completing a task, whereas extrinsic motivation reflects acceptance of the value or utility of a task. This can be conceptualized as a self-determination continuum (**Figure 1**). SDT emphasizes processes through which a person internalizes health behaviors so that they may be self-determined (Ryan et al., 2008). The theory highlights three basic human psychological needs, which when satisfied yield enhanced motivation and well-being (Ryan and Deci, 2000):

- *Autonomy:* the feeling of psychological freedom or choice.
- *Competence:* perceived self-efficacy (i.e., one's belief in one's ability to succeed).
- *Relatedness:* the need to feel belongingness and connectedness with others.

SDT has previously been employed to examine individuals' motivations for hearing aid use (Ridgeway et al., 2013, 2015), and may provide a useful framework to better understand individuals' motivations for engagement and adherence to other hearing interventions, such as CBAT. Any novel insights gained from SDT may be used to inform the future development of feasible and effective CBAT interventions for people with hearing loss.

Auditory training has the potential to be a useful intervention for people with hearing loss, including hearing aid users and those who choose not to wear hearing aids, or those who have mild hearing loss and would not necessarily benefit substantially from amplification. The present study focused on adults with

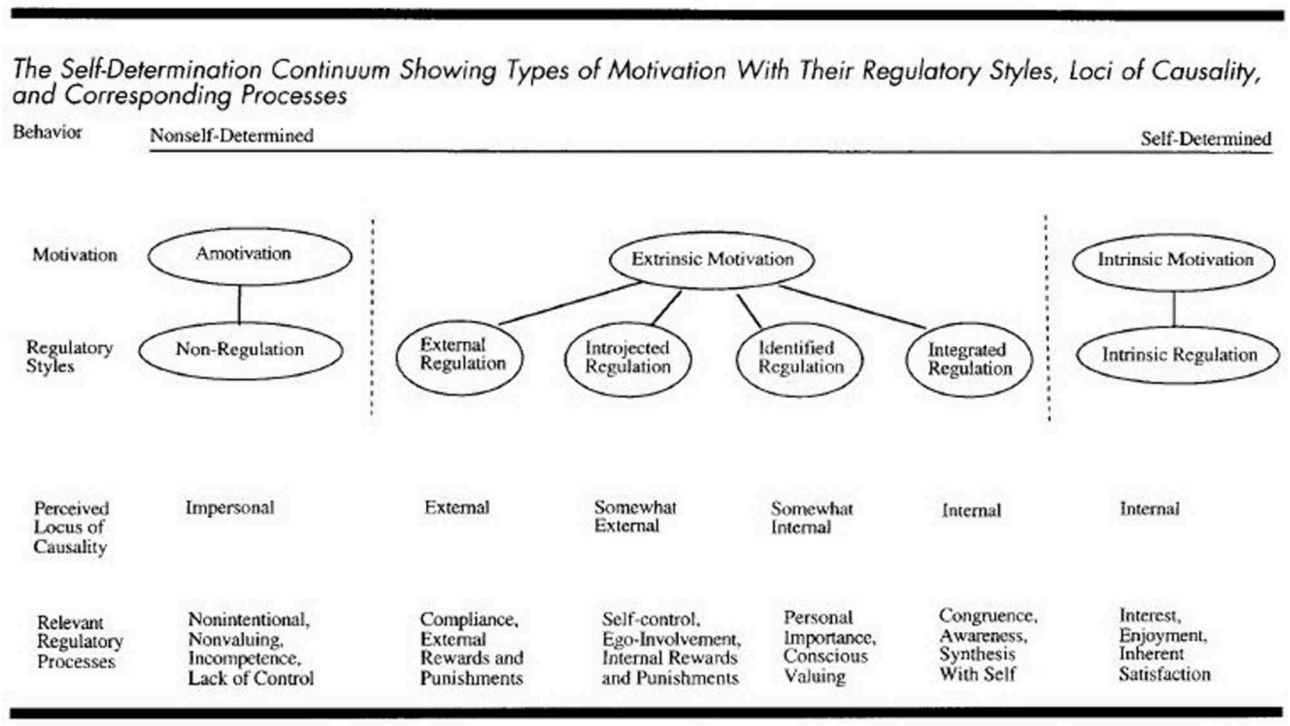


FIGURE 1 | The self-determination continuum, Ryan and Deci (2000). Copyright © 2000 by the American Psychological Association. Reproduced with permission. The official citation that should be used in referencing this material is Ryan and Deci (2000). The use of APA information does not imply endorsement by APA.

mild SNHL who were experiencing hearing difficulties, but had not yet sought intervention for their hearing loss. A randomized controlled trial (RCT) of 44 adults with mild SNHL examined the effects of a 4-week home-based program of CBAT on speech perception, cognition and self-reported hearing abilities (Ferguson et al., 2014). Participants completed a 4-week program of CBAT at home. There were high levels of adherence with CBAT, whereby 80% of participants ($n = 35$) completed the recommended amount of training (6 h over 4 weeks) and 75% ($n = 33$) exceeded the recommended training, with no participant drop outs (Ferguson et al., 2014). Findings showed significant post-training improvements in cognition and self-reported hearing abilities, particularly for challenging task conditions. However, it remains untested as to whether these benefits were readily perceived by the study participants.

The main aim of the present investigation was to explore participants' motivations for high levels of adherence, uptake and engagement with CBAT in this study using SDT as the theoretical framework. This was achieved through:

- a post-training feedback questionnaire, administered in the RCT,
- two post-RCT focus groups.

A secondary aim sought to qualitatively examine the perceived benefits of the CBAT intervention and compare this with the published (quantitative) behavioral RCT results (Ferguson et al., 2014).

Materials and Methods

The study was approved by the Nottingham Research Ethics Committee and Nottingham University Hospitals NHS Trust Research and Development. Signed, informed consent was obtained.

Participants

Randomized Controlled Trial

Adult non-hearing aid users were recruited to take part in the RCT from three General Practitioner (family physician) surgeries in Nottingham, UK (see Ferguson et al., 2014 for full details of the study design, procedure, and post-training outcomes). Participants (29 male, 15 female) were aged 50–74 years old (mean = 65.3 years, $SD = 5.7$ years) with mild, symmetrical SNHL (mean hearing thresholds averaged across 0.5, 1, 2, and 4 kHz = 32.5 dB HL, $SD = 6.0$ dB HL, with a left-right difference of <15 dB). Computer literacy ranged from "never used a computer" ($n = 7$), to "beginner" ($n = 20$), and "competent" ($n = 17$).

Focus Groups

Ten participants from the RCT (seven male, three female) volunteered to participate in one of two focus groups (five per group). Mean age was 64.8 years ($SD = 5.7$ years), and mean better ear hearing thresholds averaged across 0.5, 1, 2, and 4 kHz = 30.4 dB HL, ($SD = 6.1$ dB HL). Participants travel expenses were paid, and they received a £10 inconvenience fee for their visit.

Procedure

Randomized Controlled Trial

Participants were randomized to one of two groups in a randomized, quasi-crossover study design (Ferguson et al., 2014). The Immediate training group attended three test sessions (pre-training, post-training, and 4 week follow-up), the delayed training group attended four test sessions (control baseline, pre-training, post-training, and 4 week follow-up).

Participants completed a 4-week program of computer-based phoneme discrimination training at home, using a loan laptop which was specially programmed with only the CBAT (phoneme discrimination) training software. Training stimuli were 11 phoneme continua (/a/-/uh/, /b/-/d/, /d/-/g/, /e/-/a/, /er/-/or/, /i/-/e/, /l/-/r/, /m/-/n/, /s/-/sh/, /s/-/th/ and /v/-/w/), synthesized from end-points consisting of real voice recordings, delivered for 15 min/day, 6 days/week, for 4 weeks. The training was a 3-interval, 3-alternative forced choice task. During training, participants heard three phoneme sounds presented sequentially by three on-screen characters. They were then asked to select the character who made the "odd one out" phoneme sound. Participants completed two short (five-trial) familiarization demonstrations with the researcher in the laboratory prior to at-home training.

Training was delivered using software developed at the MRC Institute of Hearing Research (IHR-STAR) but with graphics designed for adult participants (Ferguson et al., 2014). Visual feedback (character waving) indicated correct responses to participants on a trial-by-trial basis. Participants were contacted once a week via telephone during the 4-week training period. This was to monitor participants' progress and to identify and resolve any technical or procedural issues with training.

Outcome measures were administered at each test session to assess participants' speech perception performance (Sentence and digit perception in noise tasks), cognition (single and divided attention, working memory), and self-reported hearing (Glasgow Hearing Aid Benefit Profile, Speech, Spatial and Qualities of Hearing) (Ferguson et al., 2014).

Post-training Feedback Questionnaire

At the post-training test session, a questionnaire (adapted from Benedek and Milner, 2002) was used to assess participants' views of the CBAT intervention and the usability and desirability of the training software. The questionnaire was administered to all RCT participants by interview at the post-training session and consisted of three sections:

A. *Statements:* Participants were asked to rate whether they agreed or disagreed with 10 short statements describing their experience with the CBAT software using a five-point Likert scale (strongly agree to strongly disagree).
B. *Descriptor words:* Sixty descriptor words were presented on response cards and were categorized as positive (+, $n = 35$), negative (−, $n = 23$) or could be interpreted as either positive or negative (+/−, $n = 2$). Participants identified words that described their experience with the CBAT program and of those, participants prioritized the five words most relevant to their experience.

C. *Open-ended questions:* Participants were asked three open-ended questions to assess the (i) worst, and (ii) best aspect of their experience with the training program, and (iii) any changes that would make the program more interesting, enjoyable or engaging. Content analysis (Krippendorff, 2004) was used by one researcher (HH) to develop mutually exclusive themes that identified the content of participants' responses.

Focus Groups

Three key questions were considered in the focus groups;

1. What motivated participants to take part in the CBAT study?
2. Why did participants engage and comply with the CBAT program? And,
3. What were the perceived benefits of the CBAT program?

These questions were supplemented by additional probe questions to ensure that discussions were detailed and remained on track. The focus groups lasted 2.5 and 2 h, respectively, and were each facilitated by two researchers (MAF and HH) in a quiet room, free from distraction. The majority of questions were asked by the primary facilitator (MAF).

The focus groups were audio recorded using a high quality audio recorder and transcribed verbatim. Focus group transcripts were entered into *QSR Nvivo* (Version 8). Thematic analysis was based on guidelines by Braun and Clarke (2006). To facilitate the emergence of themes, the transcripts were read, reviewed, reread and reviewed again, by one researcher (AM), to gain familiarity with the content. Analysis began with open coding to catalog what was seen to be "going on" in the data. Themes were identified by re-visiting the codes and the data, to which they had been applied, to rethink, revise and develop higher order categories.

Results

Randomized Controlled Trial

A summary of the quantitative results from the auditory training efficacy RCT are provided below. For detailed analyses, see Ferguson et al. (2014).

Auditory training: For CBAT, robust phoneme discrimination learning was found for both immediate training and delayed training groups, with the largest improvements in threshold shown for phoneme pairs with the poorest initial thresholds.
Outcome assessment: The immediate training group showed significant improvements in self-reported hearing, divided attention, and working memory. However, training did not result in consistent improvements in speech perception in noise. There was no evidence of any significant improvements in performance on any of the outcomes for the delayed training (control) group.
Follow-up assessment: Retention of benefits at 4 weeks post-training for the immediate training group was shown for phoneme discrimination, divided attention, working memory, and self-report of hearing disability.

Aim 1: Exploring Motivations for CBAT Uptake, Engagement, and Adherence

Data from the questionnaires and focus groups are interpreted as being representative of intrinsic or extrinsic motivation according to SDT (Ryan and Deci, 2000), based on the Self-Determination Continuum (**Figure 1**).

Post-training Feedback Questionnaire

Statements

Frequencies of participants' responses to the 10 statements are summarized in **Table 1**.

Intrinsic motivation

The majority of participants agreed that the CBAT intervention was both interesting and enjoyable, suggesting there was intrinsic motivation to undertake training. Most agreed or strongly agreed with the statements "The training program held my interest" ($n = 35$, 79.5%) and "I enjoyed training with the program" ($n = 38$, 86.4%), which is indicative of participants acting of their own free will (autonomy). There was little agreement shown for "I found my attention on the training program wandered during the session" ($n = 7$, 15.9%), suggesting that those participants were actively engaged with the CBAT. Finally, low agreement with the item "I found the training program difficult to use"

TABLE 1 | Number and percentage of total participants ($n = 44$) responding to statements about their experiences with the CBAT intervention. SDT = Self Determination Theory.

Statement	SDT Motivation type	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
The training program held my interest	Intrinsic	0 (0.0%)	4 (9.1%)	5 (11.4%)	23 (52.3%)	12 (27.3%)
I enjoyed training with the program	Intrinsic	0 (0.0%)	1 (2.3%)	5 (11.4%)	20 (45.5%)	18 (40.9%)
I found my attention on the training program wandered during the session	Intrinsic	16 (36.4%)	16 (36.4%)	5 (11.4%)	6 (13.6%)	1 (2.3%)
I understood what to do when using the training program	Intrinsic	0 (0.0%)	0 (0.0%)	1 (2.3%)	18 (40.9%)	24 (54.5%)
I would never use this training program again	Intrinsic	24 (54.5%)	15 (34.1%)	2 (4.5%)	3 (6.8%)	0 (0.0%)
I found the training program difficult to use	Intrinsic	27 (61.4%)	15 (34.1%)	1 (2.3%)	0 (0.0%)	1 (2.3%)
The training program did not let me know how I was doing	Intrinsic	11 (25.0%)	15 (34.1%)	4 (9.1%)	10 (22.7%)	4 (9.1%)
I felt motivated to use the training program regularly	Extrinsic	0 (0.0%)	2 (4.5%)	8 (18.2%)	21 (47.7%)	13 (29.5%)
Doing the training made me more aware of my hearing	Extrinsic	0 (0.0%)	3 (6.8%)	6 (13.6%)	19 (43.2%)	16 (36.4%)
I did the training because it might make my hearing better	Extrinsic	1 (2.3%)	7 (15.9%)	5 (11.4%)	17 (38.6%)	14 (31.8%)

($n = 1$, 2.3%) and high agreement with "I understood what to do when using the training program" ($n = 42$, 95.5%) demonstrates competence in participants' ability to undertake CBAT. Only, three participants (6.8%) agreed with the statement "I would never use this training program again."

Extrinsic motivation

The majority of participants ($n = 34$, 77.3%) agreed or strongly agreed with the statement "I felt motivated to use the training program regularly." Although the reasons for this motivation cannot be inferred from responses to this question alone, responses to item "I did the training because it might make my hearing better" ($n = 31$, 70.5%) suggested that there were extrinsic motivations for participating in the training. The majority of participants also agreed with the statement "Doing the training made me more aware of my hearing" ($n = 35$, 79.5%).

Descriptor words

Participants' descriptor word selections are presented in **Table 2**, ranked in order of frequency. Participants selected an average of 22.48 ($SD = 7.28$) words to describe their experiences with the CBAT. All descriptor words were selected at least twice across all participants. Of the first 30 words in the table, 28 are positive (93.3%) and only one is negative (3.3%), with an average of 26.5 ($SD = 12.1$) participant selections per item. Of the last 30 words in the table, seven are positive (23.3%), and 22 (73%) are negative, with an average of 6.2 ($SD = 3.4$) participant selections per item.

The five most frequently selected words "Easy to use," "Straightforward," "Organized," "Rewarding and accessible" are intrinsic in nature, and reflect autonomy and competence (i.e., participants are willing and able to complete the CBAT). Word selections such as "Valuable" and "Relevant" on the other hand suggest extrinsic motivations, whereby participants are identifying with and consciously valuing the CBAT intervention.

Frequency of selection for participants' top five prioritized descriptor words to describe their experience with the auditory training software is illustrated in a word cloud (**Figure 2**), where words with the greatest frequency of selection appear larger and darker than those words that were less frequently selected. "*Easy to use*" (intrinsic motivation) was by far the most frequently selected by participants as one of the top-five descriptors to explain their experiences with the CBAT software (28/44 participants, 63.6%). Other frequently prioritized words, selected by at least a quarter of all participants, included *Straightforward* (intrinsic; $n = 15$, 34.1%), *Valuable* (extrinsic; $n = 14$, 31.8%), *Rewarding* (extrinsic; $n = 13$, 29.5%), *Motivating* (extrinsic; $n = 12$, 27.3%) and *Useful* (extrinsic; $n = 11$, 25.0%).

Open-ended questions

1. What was the best aspect(s) of your experience with the training program?

Answers to this question were grouped into seven main themes (italicized). Themes are reported here in the order most commonly referred to, and grouped according to intrinsic and extrinsic motivations.

TABLE 2 | Frequency of participants' word choices to describe their experience with the CBAT intervention.

Rank	Descriptor word	Positive (+), negative (–) or ambiguous (+/–)	Frequency of selection by $n = 44$ participants
1	Easy to use	+	43
2	Straightforward	+	40
3	Organized	+	38
4	Rewarding	+	36
5	Accessible	+	36
6	Valuable	+	34
7	Motivating	+	34
8	Fun	+	34
9	Usable	+	32
10	Repetitive	–	31
11	Consistent	+	31
12	Useful	+	30
13	Relevant	+	30
14	Efficient	+	30
15	Simplistic	+/–	29
16	Familiar	+	26
17	Stimulating	+	25
18	Appealing	+	25
19	Reliable	+	22
20	High quality	+	20
21	Trustworthy	+	20
22	Predictable	+	19
23	Attractive	+	19
24	Comprehensive	+	18
25	Personal	+	18
26	Connected	+	18
27	Inviting	+	16
28	Fast	+	14
29	Desirable	+	13
30	Flexible	+	13
31	Fresh	+	13
32	Empowering	+	12
33	Exciting	+	11
34	Frustrating	–	10
35	Tedious	–	10
36	Collaborative	+	10
37	Boring	–	9
38	Sophisticated	+	9
39	Busy	–	9
40	Unpredictable	–	9
41	Time consuming	–	8
42	Customizable	+	8
43	Rigid	–	7
44	Unconventional	+/–	7
45	Slow	–	6
46	Confusing	–	5
47	Uncontrollable	–	5
48	Dull	–	4

(Continued)

TABLE 2 | Continued

Rank	Descriptor word	Positive (+), negative (−) or ambiguous (+/−)	Frequency of selection by $n = 44$ participants
49	Intimidating	−	4
50	Overwhelming	−	4
51	Inconsistent	−	3
52	Stressful	−	3
53	Complex	−	3
54	Gets in the way	−	3
55	Time saving	+	3
56	Hard to use	−	2
57	Not valuable	−	2
58	Overbearing	−	2
59	Patronizing	−	2
60	Too technical	−	2



FIGURE 2 | Word cloud to show participant's top five word choices describing their experience with the auditory training software.

Intrinsic motivation

An easy and enjoyable task: was reported by 12 participants (27.2%). *Sense of achievement associated with completing the training:* reported by four participants (9.0%).

Extrinsic motivation

Increased awareness of hearing or hearing difficulties: reported by eight participants (18.2%). *To aid research*: was offered as a response by one participant (2.3%).

In addition, a number of participants provided unprompted accounts of perceived benefits of the CBAT intervention, including; *improved listening, concentration, or attention post-training:* 10 participants (22.7%) and *Improved PC literacy* or *a desire to further improving their PC literacy*: reported by two participants (4.5%).

2. What was the worst aspect(s) of your experience with the CBAT program?

Responses to this question were grouped into six main themes:

Intrinsic motivation

Technical issues with training hardware or software: reported by 17 participants (38.6%). *Training tasks were repetitive or boring:* reported by seven participants (15.9%). *Feedback in the training software was not satisfactory:* five participants (11.4%) felt that the feedback did not always reflect how they perceived they were performing. *Performance on the training task:* four participants (9.1%) reported they were unhappy with their own performance in the CBAT intervention.

Extrinsic motivation

Practical issues with training: Eight participants (18.2%) reported issues such as finding time to train, or setting up and putting away a laptop computer. Finally, *Lack of experience with computers*: reported by two participants (4.5%).

3. What would you change to make the training program more interesting, enjoyable or engaging?

Responses were grouped into five main themes. All responses related to the nature of the CBAT software itself and are therefore interpreted as most relevant to intrinsic motivation.

Intrinsic motivation

Software changes: Ten participants (22.7%) reported the software could be improved, for example, changes to the feedback provided or the adaptive nature of the training games. Nine participants (20.5%) suggested *improved graphics*, seven participants (15.9%) wanted to see *changes to the sounds*, three participants (6.8%) suggested *more game-play in training tasks,* and three participants (6.8%) suggested *more variety* in the CBAT software.

Focus Groups

Thematic analysis of focus group transcripts provided main themes and sub-themes for each of the three main research questions, summarized in **Table 3**.

1. What motivated participants to take part in the auditory training study?

Participants were extrinsically motivated to take part in the CBAT study as a result of their *hearing difficulties*. Participants reported that they took part in the study either because they had noticed difficulty hearing in certain situations, or other people had commented on their hearing difficulties.

Participants talked about how their families had encouraged them to seek help for their hearing difficulties, and this prompted them to take part in the study:

> *"The family's been on at me for a long time now about me hearing and so I thought, yeah, go for it."*

Some participants had noticed their hearing difficulties themselves, and this made them to want to find out more about their hearing levels:

> *"I was concerned about my hearing, especially with the grandchildren I couldn't always hear what they were saying and I didn't want to end up like my mum."*

Some participants were not sure if their hearing was bad enough to require further attention, and reported wanting to take part in

TABLE 3 | Main categories and sub themes from thematic analysis of focus group transcripts.

Question	Main categories	Sub-themes
1. What motivated participants to take part in the CBAT study?	Hearing difficulties (extrinsic motivation)	Difficulty hearing in certain situations
		Relatives/friends commenting on their hearing difficulties
		Wanted to find out more about their hearing difficulties
		Curious about hearing loss
		A desire to improve their listening abilities
		Received invitation letter from family physician
2. Why did participants engage and comply with the CBAT program?	A sense of achievement (intrinsic motivation)	Desire to beat previous scores
		Seeing the training through to the end
	A desire and capacity to help others (extrinsic motivation)	To help others by taking part in research
		Having spare time to fill
3. What were the perceived benefits of the CBAT program?	Increased concentration, attention and focus in everyday listening	Improved listening skills
		Strategies for listening
	Encouraged further help-seeking behaviors	Seeking further hearing intervention

the research to *"catch it quick"* and to see if it would help before their hearing deteriorated further:

> *"Well I'm not sure whether I'm deaf or not. You know how you are because you're on the borderline."*

The invitation (received via their family physician) motivated some participants to take part in the study because they wanted to find out more about their hearing:

> *"I knew that I had got some impairment with my hearing anyway because I keep getting this…ay? What? I thought there's something wrong here that's not quite right and as I was thinking about that, the letter came saying, would you like to take part. I thought, that's timely, yes please."*

2. Why did participants engage with and adhere to the auditory training program?

Participants were both intrinsically and extrinsically motivated to engage in and adhere to the CBAT intervention. Intrinsic motivation to engage and adhere to the intervention was governed by participants' *sense of achievement* associated with on-task improvement and completion of the CBAT program.

Participants reported a challenging element to the training. They reported an intrinsic desire to beat their scores each session and this motivated them to continue:

> *"I was trying to beat the other score, and I thought, yeah, I'm going to get it this time."*
>
> *"Yes, every time I sat down, I wanted to beat the next one."*

Participants wanted to beat their own scores as there was a sense that if they were to improve their scores then their hearing might be improving. Therefore, an extrinsic motivation leading to adherence with training was a *desire to improve their listening abilities:*

> *"I was trying to do better every day and thinking, I'm going to get all these right, and then the first couple seemed quite easy and then it seemed to get really, really hard but it just made me want to do better, really, every time."*

Participants were intrinsically motivated by the sense of achievement gained from completing the intervention program:

> *"Oddly, like a sense of achievement, to actually complete the course, if you like."*

There was also a sense of commitment among the participants. Once they had said they would do something they wanted to see it through to the end:

> *"Well, it's the sort of thing that we've set out on a course of action, like [focus group participant] was saying. He [focus group participant] likes to finish things he has set out to do."*

A secondary theme was extrinsic in nature, *the desire and capacity to help others.* Participants commented that they completed the training because they wanted to be able to help other people with hearing difficulties. They felt that if the training worked then they might be able to recommend it to other people who might benefit from it:

> *"Another reason I started in the first place was the fact that I wanted to help others, you know…Let's go and see what this is about."*

Additionally, some participants reported a desire to aid research, and all participants reported having spare time to fill, particularly those who were retired from work:

> *"I thought well it's worth doing, it's worth looking at, seeing as I've got the time…obviously retired, and said, I've got the time to do this, let me do it now."*

Aim 2: Examining the perceived benefits of CBAT

3. What were the perceived benefits of the auditory training program?

The dominant theme was *increased concentration, attention and focus in everyday listening*. All but one participant reported that the training made them concentrate more:

> *"…It [the training] made me concentrate more, it certainly did."*

> *"I think it just made me aware that if I do want to hear what's going off, I've got to pay attention and focus more than I used to."*

Consequently the improved concentration and attention was associated with improved listening:

> *"Yes, it does make you concentrate and think—when you are concentrating you can hear more."*

Through increased concentration, participants reported post-training that they had developed better strategies for listening, such as *"looking at peoples' lips"*; *"people watching;"* and concentrating on the main conversation *"rather than trying to listen to three conversations*." This theme mirrored reports of *improved listening, concentration, or attention post-training*, offered by 10 participants in the post-training questionnaire.

A secondary theme identified in the focus groups was that training *encouraged further help-seeking behaviors*:

> "*I think the primary thing is identifying that there is a problem in the first place. We have, and so we have got the wherewithal to actually do something about it, which your program is good at…*"

Furthermore, participants thought CBAT may have the same effect on others by prompting them to seek further help:

> *"…if [after being provided with training] they think they have got a problem, that would enhance them or encourage them to go for further tests."*

After taking part in the CBAT study, two focus group participants had since been fitted with hearing aids. One of those individuals described CBAT as a stepping-stone to seeking further intervention:

> *"From the point of view of this training, I sort of looked at it as a sort of middle ground. It, I feel that it helps me, but then I subsequently needed them [hearing aids] to help me a bit more."*

Discussion

The primary aim of this investigation was to examine motivations to undertake a program of home-delivered CBAT to improve listening for adults with mild hearing loss. Self-Determination Theory (SDT; Deci and Ryan, 1985) was adopted as a theoretical framework by which to interpret motivations for initial participation in the study (uptake), engagement and adherence with the CBAT intervention using data from a post-training questionnaire and two focus groups. A secondary aim was to examine the perceived benefits of CBAT to compare with the published behavioral outcomes of this study (Ferguson et al., 2014).

Motivations for CBAT Uptake, Engagement, and Adherence

Results from the present research showed different contributions of intrinsic and extrinsic motivation for participants' uptake, and engagement and adherence, with CBAT. The main theme explaining participants' motivations for initial participation in the study (CBAT uptake) identified from the focus groups was participants' hearing difficulties (extrinsic motivation). Sub-themes included identification of hearing difficulties by relatives or friends, participants' desire to improve their listening, and the invitation into the study being received via their family physician. These results showed that participants were extrinsically motivated to take part in the study and to try CBAT in an attempt to address their hearing difficulties.

Engagement and adherence with CBAT was shown to be influenced by both intrinsic and extrinsic motivation. Descriptors of their experiences with the intervention from the post-training questionnaires was highly positive in nature, with participants' selecting positive descriptor words more frequently than negative words. Responses to the statements showed that participants found the CBAT easy to use, suggesting competence in their ability to undertake CBAT. This was the case for the vast majority of participants in the study despite a wide range of computer skills, with seven participants having never used a computer before. In addition, participants agreed that the intervention was both interesting and enjoyable. Based on SDT, this suggests that participants were intrinsically motivated to engage and comply with the intervention, and demonstrates a level of autonomy for the task. It has been argued that competence must accompany autonomy in order for individuals to see their behaviors as self-determined by intrinsic motivation (Reeve, 1996), and future CBAT programs may benefit from addressing this specifically in their design.

Results from the focus groups suggested that on a day-to-day basis, participants in the study were intrinsically motivated to adhere to training in an attempt to beat their previous scores. In the long term, participants were committed to seeing the training intervention through to the end for the inherent satisfaction associated with program completion. A secondary theme associated with CBAT adherence was the desire and capacity to help others (extrinsic motivation). Participants believed that by completing the intervention they may help other people with hearing difficulties.

Perceived Benefits of the CBAT Intervention

For the open-ended questions in the post-training questionnaire, almost a quarter of participants provided unprompted reports of improved listening, concentration, or attention post-training. Furthermore, one of the main themes explaining perceived benefits of training from thematic analysis of focus group transcripts was *increased concentration, attention and focus in everyday listening*. Focus group participants also reported that the CBAT enabled them to develop better strategies for listening, such as concentrating on the main conversation rather than trying to listen to multiple conversations at once. In the main RCT, improvements were shown for behavioral measures of complex cognition (working memory updating and divided attention), and for self-reported hearing in a group situation (Ferguson et al., 2014).

Focus group participants reported that the CBAT made them more aware of their hearing difficulties, and some participants said that taking part in the training program encouraged them to seek further intervention to address their hearing difficulties. This suggests that CBAT may act as an important stepping stone toward further intervention or help-seeking behaviors for some individuals with hearing loss.

Future Directions

"To be motivated means *to be moved* to do something" (Ryan and Deci, 2000), but evidence suggests that the maintenance and enhancement of human motivation requires supportive conditions. Although adherence was high in this study, there is evidence to suggest that adherence may be up to 50% lower in real-world clinical application (Sweetow, 2009). It is therefore important to better understand CBAT adherence in a research setting, so that CBAT interventions stand the best chance for high rates of adherence outside of the research environment.

Using SDT as a theoretical framework, we present a number of considerations for the development of future CBAT that may facilitate both intrinsic and extrinsic motivations for uptake, engagement and adherence.

Intrinsic Motivation

In the present study, when asked about the changes they would change to make the training program more interesting, enjoyable or engaging, questionnaire respondents reported a number of software developments, including increased gameplay. Gameplay has been shown to promote enjoyment and adherence with interventions in health and education domains (e.g., Nilsson et al., 2009; Papastergiou, 2009), and this has previously been examined within ENT, using gameplay to CBAT interventions for tinnitus (Hoare et al., 2014). One approach to adequately address these software considerations would be to involve the target population themselves in the design of CBAT interventions. Collaboratively involving the end-user in the development of digital and eHealth interventions that target behavior change ensures material is aligned to patient need (Ferguson, 2012), and has been shown to be critical for addressing low uptake and adherence (Kohl et al., 2013). Furthermore, this person-based approach maximizes opportunities for interventions to fully address the priorities and needs of the target population (Yardley et al., 2015).

Intrinsic motivation has previously been shown to be enhanced by positive feedback, but diminished by negative feedback (Deci, 1975). In the present study, the CBAT provided participants with trial-by-trial visual feedback for correct responses. In the open-ended questions however, five participants reported the worst aspect of their experience with the CBAT program was inconsistency in the trial-by-trial feedback they received, whereby they felt that the feedback did not reflect how they perceived they were performing. As such, it is possible that this may have affected their levels of intrinsic motivation. Ensuring consistency in the trial-by-trial feedback is therefore a key consideration for future CBAT development that aims to maximize intrinsic motivation.

Extrinsic Motivation

The main theme contributing to participation in this study was participants' hearing difficulties. This provides a direct link between participants' health condition and their motivations for taking up the CBAT intervention. Furthermore, participants reported being influenced by the fact that the invitation to join the study was received via their family physician. In a recent examination of patient perceptions of benefit and enjoyment of auditory training, Tye-Murray and colleagues suggest that compliance with auditory training might be further enhanced if patients have regular contact with a hearing professional and train with meaning-based materials (Tye-Murray et al., 2012).

A substantial body of research has demonstrated that contexts that are supportive of autonomy, competence, and relatedness foster greater internalization and integration of behaviors, and therefore facilitate extrinsic motivation (Ryan and Deci, 2000). With this in mind, a number of recommendations can be made to increase extrinsic motivation in future CBAT design.

Autonomy

Within SDT, extrinsic motivation can demonstrate different levels of internalization. For example, individuals may either personally grasp that a CBAT intervention may offer benefit to address their hearing difficulties, and subsequently they adhere to the agreed intervention through personal endorsement (high internalization). Alternatively, individuals can be recommended or instructed to do the training, signaling compliance with external guidance (low internalization). In order to internalize behaviors, individuals must be able to relate to that meaning in terms of their own goals and values. Autonomy refers to choice and freedom from external pressure to behave in a certain way. Providing individuals with the freedom of task selection within CBAT interventions (i.e., the choice to select tasks that are most relevant to them and their hearing difficulties) may therefore help promote the personal importance of the intervention and increase conscious valuing (Tye-Murray et al., 2012).

Competence

People are more likely to adopt activities that they feel they can be effective in. So, any future CBAT development should bear this in

mind. Ensuring that CBAT software is easy to use and achievable will help achieve and maintain usability and desirability.

Relatedness

Behaviors are reinforced when they are prompted, modeled, or valued by significant others. Within hearing rehabilitation there is growing evidence to suggest that the involvement of significant others offers additional benefit to individuals with hearing loss (Preminger, 2003; Pichora-Fuller et al., 2013). Furthermore, the significant other themselves may also gain benefit from their involvement (Pyykkö et al., 2014). Thus, the involvement of significant others in the delivery and ongoing support of CBAT interventions may serve to promote relatedness and increase motivation.

Limitations

It should be noted that one of the researchers (HH) was involved in quantitative data collection for outcome measures at some participant tests sessions, and was also present in the two post-training focus groups. Although HH was not the primary facilitator, we are unable to rule out the possibility that the involvement of HH in both quantitative and qualitative data collection may have influenced the qualitative data for some participants.

A secondary theme from the focus groups accounting for engagement and adherence with CBAT was participants' desire to help other people with hearing loss by taking part in research. This was particularly true for those participants who had retired from work and had time to spare. As such, this is unlikely to be a factor associated with engagement and adherence with CBAT outside of a research environment. Participants who took part in the RCT and subsequent focus groups were a volunteer sample, and therefore may have different motivations than people with hearing loss who did not choose to take part. In addition, the nature of qualitative research means that findings cannot be universally applied to other populations. This study involved people with mild hearing loss who did not have hearing aids. It is possible that people with greater degrees of hearing loss or those who had already received an intervention (e.g., hearing aid users) may have different motivations for CBAT uptake, engagement and adherence.

Although informative in the short-term, the results of this research do not provide information about engagement and adherence to CBAT over time. One of the main intrinsic motivation factors associated with engagement and adherence in this study was that the training task was simple and easy to use. However, it is not clear from this investigation whether the simplistic nature of the task could lead to boredom or frustration over an extended period. It is also important to note that the training schedule in this study was substantially shorter than other CBAT programs such as LACE, and this may have contributed to the positive appraisal of the intervention, and to the high adherence rates witnessed in this study.

Participants in the RCT received weekly phone calls to identify any technical or procedural issues with the training software, which may or may not have contributed to the high training adherence witnessed in this study. Nevertheless, the telephone calls were not reported by participants to be a factor associated with training in either the post-training questionnaire responses or the focus groups.

As is the case with all research, the participants who took part in this investigation were volunteers. As such, it should be acknowledged that these individuals might be more motivated to engage with and adhere to the CBAT intervention than individuals in the general population. Furthermore, we cannot firmly rule out the effects of individuals' expectations regarding the benefits of the CBAT intervention. Finally, due to the high rates of adherence with CBAT witnessed in the RCT, the findings of this investigation cannot inform us about how we might best support adherence for individuals who may be less motivated to adhere.

Summary and Conclusions

Self-management of hearing loss requires motivation and dedication. Participants in this study readily perceived benefits of a 4-week CBAT intervention in terms of improved concentration and attention. Participants reported that the CBAT also made them more aware of their hearing difficulties, and prompted some individuals to seek further intervention (hearing aids) to address these difficulties.

Initial participation in the study (CBAT uptake) was associated with extrinsic motivation arising from participants' hearing difficulties, whereas engagement and adherence with CBAT was influenced by both intrinsic and extrinsic motivation including a desire to beat previous scores on the training task, and to help others with hearing loss.

The use of SDT as a theoretical framework to retrospectively interpret data in this investigation has provided useful insights into the applied nature of human motivation for CBAT. Furthermore, this approach offers a framework from which to develop future research to explicitly assess individuals' motivations for audiological intervention that maximize their intrinsic and extrinsic motivations for adherence. We see this as an important first-step in informing future theory-driven development of CBAT.

Author Contributions

MF designed the study. HH, AM, and MF analyzed and interpreted the data. HH and AM wrote the manuscript. MF and HH contributed to critical discussions. HH revised the manuscript. All authors approved the final version of the manuscript for publication. All authors agree to be accountable for all aspects of the work and in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgments

The authors would like to thank Daniel Clark and Alison Riley for their assistance with the questionnaire data collection, and all

study and focus group participants for their involvement in the research. The authors would also like to thank Ariane Laplante-Lévesque for her valuable comments on an earlier version of this manuscript. Early data from the focus groups reported in this manuscript were presented at a British Academy of Audiology annual conference, abstract by Henshaw et al. (2012).

The data and analyses presented in this manuscript have not been previously published. This paper presents independent research funded by the National Institute for Health Research (NIHR) Biomedical Research Unit Programme. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References

Action on Hearing Loss. (2011). *Hearing Matters*. London: Action on Hearing Loss

Amitay, S., Irwin, A., and Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nat. Neurosci.* 9, 1446–1448. doi: 10.1038/nn1787

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4357–4362. doi: 10.1073/pnas.1213555110

Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, M. K. (2009). The emergence of Cognitive Hearing Science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x

Benedek, J., and Milner, T. (2002). "Measuring desirability: new methods for evaluating desirability in a usability lab setting," in *Proceedings of the Usability Professional Association Conference* (Orlando, FL).

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Chisolm, T. H., Saunders, G. H., Frederick, M. T., McArdle, R. A., Smith, S. L., and Wilson, R. H. (2013). Learning to listen again: the role of compliance in auditory training for adults with hearing loss. *Am. J. Audiol.* 22, 339–342. doi: 10.1044/1059-0889(2013/12-0081)

Davis, A., Smith, P., Ferguson, M., Stephens, D., and Gianopoulos, I. (2007). Acceptability, benefit and costs of early screening for hearing disability: a study of potential screening tests and models. *Health Technol. Assess.* 11, 1–294. doi: 10.3310/hta11420

Deci, E. L. (1975). *Intrinsic Motivation.* New York, NY: Plenum. doi: 10.1007/978-1-4613-4446-9

Deci, E., and Ryan, R. (1985). *Intrinsic motivation and self-regulation in human behavior*. New York, NY: PlenumPress. doi: 10.1007/978-1-4899-2271-7

Ferguson, M. (2012). Delivery of hearing healthcare and education using the internet. *ENT Audiol. News* 21, 68–69.

Ferguson, M., and Henshaw, H. (2015a). Computer and internet interventions to optimise listening and learning for people with hearing loss: accessibility, use and adherence. *Am. J. Audiol.* doi: 10.1044/2015_AJA-14-0090. (in press).

Ferguson, M. A., and Henshaw, H. (2015b). Auditory training can improve working memory, attention and communication in adverse conditions for adults with hearing loss. *Front. Psychol.* 6:556. doi: 10.3389/fpsyg.2015.00556

Ferguson, M. A., Henshaw, H., Clark, D., and Moore, D. (2014). Benefits of phoneme discrimination training in a randomized controlled trial of 50–74 year olds with mild hearing loss. *Ear Hear.* 35, e110–e121. doi: 10.1097/AUD.0000000000000020

Heinrich, A., Henshaw, H., and Ferguson, M. A. (2015). The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front. Psychol.* 6:782. doi: 10.3389/fpsyg.2015.00782

Henshaw, H., and Ferguson, M. A. (2013). Efficacy of individual computer-based auditory training for people with hearing loss: a systematic review of the evidence. *PLoS ONE* 8:e62836. doi: 10.1371/journal.pone.0062836

Henshaw, H., and Ferguson, M. A. (2014). "Assessing the benefits of auditory training to real-world listening: identifying appropriate and sensitive outcomes," in *Proceedings of ISAAR 2013: Auditory Plasticity – Listening with the Brain. 4th Symposium on Auditory and Audiological Research*, eds T. Dau, S. Santurette, J. C. Dalsgaard, L. Trangjaerg, T. Andersen and T. Poulsen (Nyborg: The Danavox Jubilee Foundation), 45–52.

Henshaw, H., McCormack, A., and Ferguson, M. A. (2012). Auditory training: exploring participant motivations, engagement and compliance. *Int. J. Audiol.* 51, 263–264.

Hoare, D. J., van Labeke, N., McCormack, A., Sereda, M., Smith, S., Al Taher, H., et al. (2014). Gameplay as a source of intrinsic motivation in a randomized controlled trial of auditory training for tinnitus. *PLoS ONE* 9:e107430. doi: 10.1371/journal.pone.0107430

Kohl, L. F., Crutzen, R., and De Vries, N. K. (2013). Online prevention aimed at lifestyle behaviors: a systematic review of reviews. *J. Med. Internet Res.* 15, e146. doi: 10.2196/jmir.2665

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., and Resnick, S. M. (2011a). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology* 25, 763–770. doi: 10.1037/a0024238

Lin, F. R., Metter, E. J., O'brien, R. J., Resnick, S. M., Zonderman, A. B., and Ferrucci, L. (2011b). Hearing loss and incident dementia. *Arch. Neurol.* 68, 214–220. doi: 10.1001/archneurol.2010.362

McCormack, A., and Fortnum, H. (2013). Why do people fitted with hearing aids not wear them? *Int. J. Audiol.* 52, 360–368. doi: 10.3109/14992027.2013.769066

Nilsson, C. N., Serafin, S., and Nordahl, R. (2009). Gameplay as a source of intrinsic motivation for individuals in need of ankle training or rehabilitation. *Presence* 21, 69–84. doi: 10.1162/PRES_a_00090

Pagoto, S., and Bennett, G. G. (2013). How behavioral science can advance digital health. *Transl. Behav. Med.* 3, 271–276. doi: 10.1007/s13142-013-0234-z

Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: a literature review. *Comput. Educ.* 53, 603–622. doi: 10.1016/j.compedu.2009.04.001

Pichora-Fuller, M. K., Dupuis, K., Reed, M., and Lemke, U. (2013). Helping older people with cognitive decline communicate: hearing aids as part of a broader rehabilitation approach. *Semin. Hear.* 34, 308–330. doi: 10.1055/s-0033-1356643

Pichora-Fuller, M. K., and Levitt, H. (2012). Speech comprehension training and auditory and cognitive processing in older adults. *Am. J. Audiol.* 21, 351–357. doi: 10.1044/1059-0889(2012/12-0025)

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Preminger, J. E. (2003). Should significant others be encouraged to join adult group audiologic rehabilitation classes? *J. Am. Acad. Audiol.* 14, 545–555. doi: 10.3766/jaaa.14.10.3

Pyykkö, I., Manchaiah, V. K., Kentala, E., and Levo, H. (2014). Significant others of patients with hearing and balance disorders report positive experiences. *Int. J. Audiol.* 53, 285–286. doi: 10.3109/14992027.2013.869840

Reeve, J. (1996). *Motivating Others.* Needham Heights, MA: Allyn and Bacon.

Ridgeway, J., Hickson, L., and Lind, C. (2013). Self-Determination Theory: Motivation and hearing aid adoption. *J. Acad. Rehabil. Audiol.* XLVI, 11–37.

Ridgeway, J., Hickson, L., and Lind, C. (2015). Autonomous motivation is associated with hearing aid adoption. *Int. J. Audiol.* 54, 478–484. doi: 10.3109/14992027.2015.1007213

Rönnberg, J., Rudner, M., and Lunner, T. (2011). Cognitive hearing science: the legacy of stuart gatehouse. *Trends Amplif.* 15, 140–148. doi: 10.1177/1084713811409762

Ryan, R. M., and Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp. Educ. Psychol.* 25, 54–67. doi: 10.1006/ceps.1999.1020

Ryan, R. M., Patrick, H., Deci, E. L., and Williams, G. C. (2008). Facilitating health behaviour change and its maintenance: interventions based on Self-Determination Theory. *Eur. Health Psychol.* 10, 2–5.

Schow, R., and Nerbonne, M. (2006). *Introduction to Audiologic Rehabilitation.* Boston, MA: Pearson Education.

Schroeder, S. A. (2007). Shattuck lecture. We can do better–improving the health of the American people. *New Engl. J. Med.* 357, 1221–1228. doi: 10.1056/NEJMsa073350

Sweetow, R. (2009). "Integrating LACE into a busy clinical practice," in *American Academy of Audiology,* (Dallas, TX).

Sweetow, R. W., and Henderson Sabes, J. (2006). The need for and development of an adaptive Listening and Communication Enhancement (LACE) program. *J. Am. Acad. Audiol.* 17, 538–558. doi: 10.3766/jaaa.17.8.2

Sweetow, R. W., and Henderson Sabes, J. (2010). Auditory training and challenges associated with participation and compliance. *J. Am. Acad. Audiol.* 21, 586–593. doi: 10.3766/jaaa.21.9.4

Tye-Murray, N., Sommers, M. S., Mauze, E., Schroy, C., Barcroft, J., and Spehar, B. (2012). Using patient perceptions of relative benefit and enjoyment to assess auditory training. *J. Am. Acad. Audiol.* 23, 623–634. doi: 10.3766/jaaa.23.8.7

Vermeire, E., Hearnshaw, H., van Royen, P., and Denekens, J. (2001). Patient adherence to treatment: three decades of research. A comprehensive review. *J. Clin. Pharm. Therapeut.* 26, 331–342. doi: 10.1046/j.1365-2710.2001.00363.x

Whitmer, W. M., Howell, P., and Akeroyd, M. A. (2014). Proposed norms for the Glasgow hearing-aid benefit profile (Ghabp) questionnaire. *Int. J. Audiol.* 53, 345–351. doi: 10.3109/14992027.2013.876110

World Health Organization. (2003). *Adherence to Long-term Therapies: Evidence for Action*. Geneva: WHO Press.

World Health Organization. (2008). *The Global Burden of Disease: 2004 Update*. Geneva: WHO Press.

Yardley, L., Morrison, L., Bradbury, K., and Muller, I. (2015). The person-based apporach to intervention development: application to digital health-related behaviour change interventions. *J. Med. Internet Res.* 17, e30. doi: 10.2196/jmir.4055

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Henshaw, McCormack and Ferguson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 11 August 2015
doi: 10.3389/fpsyg.2015.01065

Subjective ratings of masker disturbance during the perception of native and non-native speech

Lisa Kilman [1,2], Adriana A. Zekveld [1,2,3], Mathias Hällgren [2,4] and Jerker Rönnberg [1,2]*

[1] *Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden,* [2] *Linnaeus Centre HEAD, The Swedish Institute for Disability Research, Linköping University and Örebro University, Linköping, Sweden,* [3] *ENT/Audiology and EMGO+ Institute for Health and Care Research, VU University Medical Center, Amsterdam, Netherlands,* [4] *Department of Otorhinolaryngology, Section of Audiology, Linköping University Hospital, Linköping, Sweden*

The aim of the present study was to address how 43 normal-hearing (NH) and hearing-impaired (HI) listeners subjectively experienced the disturbance generated by four masker conditions (i.e., stationary noise, fluctuating noise, Swedish two-talker babble and English two-talker babble) while listening to speech in two target languages, i.e., Swedish (native) or English (non-native). The participants were asked to evaluate their noise-disturbance experience on a continuous scale from 0 to 10 immediately after having performed each listening condition. The data demonstrated a three-way interaction effect between target language, masker condition, and group (HI versus NH). The HI listeners experienced the Swedish-babble masker as significantly more disturbing for the native target language (Swedish) than for the non-native language (English). Additionally, this masker was significantly more disturbing than each of the other masker types during the perception of Swedish target speech. The NH listeners, on the other hand, indicated that the Swedish speech-masker was more disturbing than the stationary and the fluctuating noise-maskers for the perception of English target speech. The NH listeners perceived more disturbance from the speech maskers than the noise maskers. The HI listeners did not perceive the speech maskers as generally more disturbing than the noise maskers. However, they had particular difficulty with the perception of native speech masked by native babble, a common condition in daily-life listening conditions. These results suggest that the characteristics of the different maskers applied in the current study seem to affect the perceived disturbance differently in HI and NH listeners. There was no general difference in the perceived disturbance across conditions between the HI listeners and the NH listeners.

Keywords: perceived disturbance, native, non-native, speech maskers, noise maskers, working memory

OPEN ACCESS

Edited by:
Robert J. Zatorre, McGill University, Canada

Reviewed by:
Michel Hoen, Oticon Medical, France
Piia Astikainen, University of Jyväskylä, Finland

****Correspondence:***
Lisa Kilman, Department of Behavioral Sciences and Learning, Linköping University, Mäster Mattias Väg, S-581 83 Linköping, Sweden
lisa.kilman@liu.se

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *10 March 2015*
Accepted: *13 July 2015*
Published: *11 August 2015*

Citation:
Kilman L, Zekveld AA, Hällgren M and Rönnberg J (2015) Subjective ratings of masker disturbance during the perception of native and non-native speech. Front. Psychol. 6:1065. doi: 10.3389/fpsyg.2015.01065

Introduction

Listening in noisy environments can be strenuous for one and all. Even so, people seem to differ in their subjective evaluation of the impact of disturbing sounds on speech perception. This may be due to a variety of factors and knowledge of these factors provides insight into how individuals experience listening in challenging situations. One relevant individual factor is hearing acuity, i.e., whether the individual is normal-hearing (NH) or hearing-impaired (HI). Individuals with

hearing loss are more likely to have difficulties in difficult listening situations than NH individuals (McCoy et al., 2005; Tun et al., 2009; Zekveld et al., 2010, 2011). Other aspects that might affect the outcome are age and cognitive functions, as well as the characteristics of the target and the background maskers.

In this study we evaluate how NH and HI listeners perceive disturbance of different types of maskers (stationary, fluctuating, babble Swedish, and babble English) in native and non-native languages.

Previous research indicates that some types of background maskers are considered more challenging than others (Pichora-Fuller, 2009). For example, speech perception in fluctuating maskers is experienced more demanding than listening to speech in stationary maskers (Pichora-Fuller, 2009). It is also known that HI listeners have more difficulties to listen "in the dips" that exist in fluctuating maskers than NH listeners (Festen and Plomp, 1990; Versfeld and Dreschler, 2002). Human speech though, appears to have a special position as a background sound, in particular when it is intelligible. In fact, subjective ratings of perceived disturbance have been found to be associated with the intelligibility of ambient speech maskers; the higher the intelligibility, the higher the disturbance ratings (Venetjoki et al., 2006). However, in objective measures of performance, several studies have confirmed that when the background speech consists of an unfamiliar language or less well mastered language, the result is usually a release in masking (Rhebergen et al., 2005; Van Engen and Bradlow, 2007; Calandruccio et al., 2010; Van Engen, 2010; Gautreau et al., 2013; Kilman et al., 2014). Furthermore, in the study of Calandruccio et al. (2013), when the background speech consisted of linguistically and phonetically distant (English target and Mandarin masker) versus close (English target and Dutch masker) language pairs; the listener performance increased when the distance increased. In this study, we do not measure objective performance but subjective ratings and it is not for certain that performance and perceived disturbance reflect matching result.

Perceived disturbance is influenced by several factors: it is partly based on difficulties in separating similar signals (Brouwer et al., 2012) and partly on the meaningful content of the distracting speech (e.g., Pichora-Fuller, 2009). Speech in background maskers might also be perceived differently for HI listeners as compared to the listeners, due to the hearing impairment *per se*. Moore (1985) argued that impaired temporal and spectral resolution is a key factor explaining the difficulties experienced by HI individuals to understand speech in background sounds.

It has been suggested that persons with hearing-impairment have to invest more processing resources to recognize spoken words than individuals with NH (Rabbitt, 1991).It is likely that this additional investment may contribute to the fatigue experienced by HI individuals at the end of the day. Research regarding this topic shows that individuals with hearing loss need more time after work to rest and recovery (Nachtegaal et al., 2009).

When an individual is focusing on a conversation and this conversation is disturbed by competing sound, it is plausible that the attention of the individual is captured by the interfering sound (Mattys et al., 2012). Yet, it is also plausible that the individual tries to re-focus his/her attention on the conversation. However, this may require a "cost" associated with dividing attention and separating the sound and the target signals (Mattys et al., 2012). Such processing could increase the level of attentional effort, i.e., the effort it takes to ignore the distracter and selectively attend to the target (Mattys et al., 2012; Koelewijn et al., 2014).

In the current study, the aim was to assess perceived *disturbance* from a masker during speech perception. We suggest an association between perceived *disturbance* and perceived *effort*. Effort is here assumed to be a consequence of perceived disturbance. Listening *effort* has been defined as "the mental exertion required to attend to and understand an auditory message" (McGarrigle et al., 2014). Listening may become effortful as a result of background noise, hearing impairment (McGarrigle et al., 2014) and/or being a non-native speaker of the target language (Mattys et al., 2012). The definition of "*disturbance*" is "the interruption of a settled and peaceful condition" (Oxford English Dictionary). In the context of the current study (i.e., speech perception) the definition of disturbance is: "The interruption of intended listening." As a result, the attentional focus may be directed toward the interrupting sound. It has been claimed that the degree of auditory disturbance, i.e., the ability to control attention and avoid distraction, can be attributed to individual differences in working memory capacity (Conway et al., 2001; Kane et al., 2001; Sörqvist et al., 2012). High working memory capacity individuals seem to have more steadfast focus of attention and less processing of the background sound (Sörqvist and Rönnberg, 2014).

The relationship between working memory and language understanding is explained in the framework of ease of language understanding (ELU; Rönnberg, 2003; Rönnberg et al., 2008, 2013). Generally, the model clarifies the relationship between implicit and explicit functions during language processing. Furthermore, the mismatch function in the model explains the concept of perceived disturbance. When the listening situation is relatively undisturbed, the incoming semantic signal can be matched to the stored language representations in long-term memory. In that case, lexical access proceeds implicitly with ease, and language understanding is established. However, if the language signal is degraded by noise, hearing impairment and/or a non-native language, a mismatch may occur and the listener will have difficulties understanding the message. The more degraded the signal is, the more likely that the listener will experience the mismatch as more disturbing. Or expressed differently: The degree of mismatch outlines the degree of perceived disturbance. Additionally, for degraded speech, listeners will have difficulties to find language representations in the long-term memory and will as a consequence have to employ explicit processing in an attempt to comprehend the message. Thus, working memory must be invoked in order to succeed in language understanding. The ELU model describes that the degree of listening effort is related to the amount of explicit cognitive resources required to disentangle the fuzziness between the language input and the stored language representations in the long-term memory.

Even though listening in noise and its negative consequences are well documented (e.g., Kjellberg et al., 1996; Larsby et al., 2005; Jahncke and Halin, 2012; Hua et al., 2013), the main focus in studies applying subjective noise- and disturbance-ratings is usually the impact of environmental sounds. For example, the

disturbance of office noise and traffic/railway/aircraft noise is commonly assessed. Furthermore, previous studies within the field of speech perception have focused on listening effort and how it can be measured objectively (Kramer et al., 1997; Murphy et al., 2000; Tun et al., 2009; Zekveld et al., 2010) and subjectively (Larsby et al., 2005; Zekveld et al., 2010). Studies in speech perception measuring self-rated disturbance are sparse and have mainly focused on simulated workplace-settings, like office noise, daycare and traffic settings (Hua et al., 2014), or perceived effort and disturbance when completing a task in office noise (Hua et al., 2013). To our knowledge, there is currently no empirical study of subjectively rated masker disturbance during the perception of masked native and non-native speech.

In the present study, we therefore evaluated the perceived disturbance for NH and HI listeners perceiving Swedish and English target speech in different masker conditions, including stationary and fluctuating noise and two-talker babble in Swedish and English. The subjective ratings analyzed in the present study were collected in the context of a larger study (Kilman et al., 2014, 2015).

Hearing impairment is commonly associated with increased listening disturbance (Hua et al., 2014; Skagerstrand et al., 2014).Therefore, we hypothesized that HI listeners will experience the different speech and masker conditions as generally more disturbing than the NH listeners.

Speech is generally considered to be more interfering than other sound sources (Venetjoki et al., 2006). Consequently, we hypothesized that both NH listeners and HI listeners will rate the speech maskers as more disturbing than the two noise-maskers in both target languages. Interactions between target language, masker conditions and hearing status were expected, but there is no firm theoretical basis for the exact pattern of disturbance.

Materials and Methods

Participants

Forty-three participants; 22NH (12 females and 10 males) and 21 HI (12 females and 9 males) were recruited for the study. In the NH group, the ages ranged from 28 to 64 years ($M = 49.5$, SD = 9.8) and in the HI group, the ages ranged from 28 to 65 years ($M = 50.1$, SD = 10.2). There was no significant difference in age between the NH group and the HI group [t (41) = 0.25, $p = 0.804$]. The NH participants were recruited from workplaces in Linköping and the HI from the audiology clinic at Linköping University Hospital, Sweden. In the NH group, education ranged from 11 to 21 years ($M = 15.8$) and in the HI group, education ranged from 8 to 21.5 years ($M = 13.7$). There was a significant difference in education between the NH group and the HI group [t (40) = −2.15, $p < 0.05$].

All participants were native Swedish speakers and had learned English as NH children in the Swedish school-system. Additional inclusion criteria for the HI participants were that they had an acquired bilateral, symmetrical sensorineural hearing loss with no severe tinnitus complaints. The study was approved by the regional ethics committee in Linköping and all participants provided written informed consent. All testing took place at Linköping university hospital and the participants received a small gift for taking part in the study.

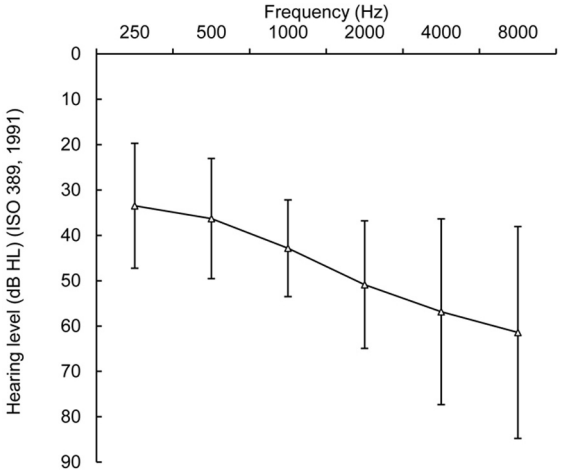


FIGURE 1 | Means and SDs (error bars) of the unaided pure-tone audiometric thresholds of the hearing-impaired participants. Hearing thresholds are averaged over both ears.

Stimuli and Tests

Pure Tone Audiometry

Pure-tone average thresholds of the NH and HI participants at the frequencies 500, 1000, 2000, and 4000 Hz were measured in the beginning of the test session. The NH participants had pure tone hearing thresholds of a maximum of 20 dB HL between 250 and 2000 Hz and a maximum of 35 dB HL at 4000 Hz. One participant had a threshold of 45 dB HL at 4000 Hz in one ear. For the HI participants, the average hearing threshold across frequencies (PTA_4) was 46.7 dB HL (SD = 10.7 dB HL). The PTA_4 ranged from 25.0 dB HL to 71.3 dB HL. The average degree of hearing loss varied from slight (16–25 dB; $n = 1$) through mild (26–40 dB; $n = 6$), moderate (41–55 dB; $n = 11$), moderately severe (56–70 dB; $n = 2$) to severe (71–90 dB; $n = 1$) (Clark, 1981) (**Figure 1**).

SRT in Noise and Speech

The SRT test was used to measure sentence intelligibility (Plomp and Mimpen, 1979) in Swedish (Hällgren et al., 2006) and in American English HINT (Nilsson et al., 1994). The HINT sentences are short and ordinary, phonemically balanced and grouped in 25 lists with 10 sentences in each. The HINT sentences were recorded with a male native speaker in Swedish and a male native speaker in English. Eight conditions were employed; two target language conditions, Swedish and English and four masker conditions; stationary masker, fluctuating masker, two-talker babble Swedish and two-talker babble English (see description below). Every condition consisted of 20 sentences and the conditions were counterbalanced across the participants. Every sentence was used only once. The masker onset occurred 3 s before speech onset and masker off-set was 1 s after speech off-set. For the NH participants, the speech was presented at a fixed level of 65 dB SPL. For the HI participants, the presentation levels of

the target speech and masker were off-line adapted according to the Cambridge prescription formula (Moore and Glasberg, 1998) based on pure tone thresholds of the best ear. A stepwise two-up-two-down adaptive procedure (Plomp and Mimpen, 1979) was to determine the level of the masker for each sentence, targeting an SNR required to perceive 50% of the sentences correctly.

The stationary masker was a speech-shaped noise developed by Nilsson et al. (1994) and by Hällgren et al. (2006).

The fluctuating masker was created from the speech-shaped noise of the target language with the same envelope fluctuations as the two-talker babble in Swedish or English (depending on the target language). The envelopes were extracted by applying a low-pass filter with cut-off frequency of 32 Hz (for details see Agus et al., 2009). Two fluctuating maskers were used, one was matched spectrally to the Swedish target and temporally to the Swedish babble and one was matched spectrally to the English target and temporally to the English babble.

Two-talker Babble Maskers The Swedish two-talker babble was created by mixing the soundtracks from a native female and a native male reading Swedish newspapers. The English two-talker babble was created by mixing the soundtracks from one native British English male and one American English female reading English/American newspapers.

Subjective Ratings

The participants were instructed to rate the perceived listening disturbance immediately after completing each condition. The participants were given a sheet of paper with eight scales, one for each condition and were asked to answer the following question: "*How disturbing was the noise you just heard?*" The question was the headline on the paper. The disturbance rating scales ranged from 0 to 10 on a continuous scale, where 10 represented "extremely disturbing" and 0 "not disturbing at all."

Results

The means and standard deviations of the perceived disturbance in the eight different SRT conditions are shown in **Table 1**. The most disturbing masker for the HI listeners seems to be the *babble Swedish in the Swedish target language*. The most disturbing masker for the NH listeners seems to be the *babble Swedish in the English target language*.

Analysis of variance (ANOVA) was conducted to assess the *impact of the two target languages (Swedish and English) and the four masker types (stationary noise, fluctuating noise, babble Swedish and babble English)* as within participant factors on the perceived disturbance for HI listeners and NH listeners (i.e., the between-participant factor). The ANOVA showed a main effect of *masker type*; F (3,123) = 5.4, $p < 0.05$, eta squared = 0.12, suggesting a moderate to large effect, but no main effect of hearing status. Also, a *significant three-way-interaction between group, language and masker type* was observed; F (3,123) = 6.53, $p < 0.001$, eta squared = 0.14, suggesting a large effect. The result indicates that the interaction effect between target language and masker type differed between the NH listeners and the HI listeners, as generally expected. Follow-up analysis of simple effects showed that there was a significant interaction between *target language* and *masker type* for the HI listeners; F (3, 60) = 6.8, $p < 0.001$, $d = 0.25$, suggesting a small significance (For the calculation of d from dependent t-test, we used the formula described in Dunlap et al., 1996, s 171). There was no significant interaction for the NH listeners; F (3, 63) = 1.6, $p = 0.19$. This result reflects that for HI listeners, there was a difference in perceived disturbance *between the maskers for the Swedish and English target languages*. No significant effects were found of target language; F (1, 41) = 1.64, $p = 0.13$, or group, as between-participant factor; F (1, 38) = 3.7, $p = 0.06$.

We expected the *speech maskers (Swedish and English babble in both target languages)* to be perceived more disturbing than the *noise maskers (stationary and fluctuating maskers in both target languages)*. We tested whether this was the case separately for the NH listeners and HI listeners. For the NH listeners, the *speech maskers* were perceived as more disturbing than the *noise maskers*; t (21) = 2.57, $p < 0.05$, $d = 0.34$, suggesting a small to moderate significance. However, for the HI listeners, the *speech maskers* were not perceived as more disturbing than the *noise maskers*; t (20) = 1.65, $p = 0.114$.

HI Listeners Probing the overall three-way interaction further, a *post hoc*, pairwise comparison (Bonferroni adjusted for multiple comparison at the 0.05 level) of the differences in disturbance ratings between the masker types across languages confirmed a significant difference for the HI listeners for the *Swedish babble*, *between* the *Swedish* and the *English target languages*; t (20) = 4.70, $p < 0.001$, $d = 0.81$, suggesting a large significance. This demonstrates that the perceived disturbance for the HI listeners in the *Swedish babble* was larger for *Swedish* as compared to the *English target language*. None of the differences in perceived disturbance of the other masker types (i.e., *stationary noise*, *fluctuating noise*, and *babble English*) between the two target languages were statistically significant; t (20) = −1.20 to −1.76, all $p > 0.05$.

Significant differences (Bonferroni adjusted for multiple comparison at the 0.05 level) were shown between the *Swedish babble* and each of the other maskers (*stationary*, *fluctuating*, *English babble*) for the *Swedish target language*, t (20) = 2.7–3.9, $p < 0.05$, $d = 0.93$ (SweBS/SweSt), $d = 0.73$ (SweBS/SweFl), $d = 0.60$ (SweBS/SweBE), suggesting a moderate to large significance for the differences. The result indicated that the HI listeners perceived *the Swedish babble* as more disturbing than the other three maskers in *Swedish target language*. No significant differences were found between the maskers for the *English target speech*; t (20) = 1.45–2.17, all $p > 0.05$ (**Figure 2**).

NH Listeners The same *post hoc* pair-wise comparisons were performed on the data of the NH listeners (independent t-tests with Bonferroni adjustment for multiple comparison at the 0.05 level). There were no significant differences in perceived disturbance from the maskers between the two *target languages* for the NH listeners. For the *English target language*, the results show significant differences between the *stationary masker* and the *Swedish babble*, t (21) = 3.5, $p < 0.05$, $d = 0.62$, suggesting a moderate significance, and between the *fluctuating masker* and the *Swedish babble*, t (21) = 3.0, $p < 0.05$, $d = 0.50$, suggesting a moderate significance. These results indicate that the perceived disturbance of the *Swedish babble* was larger than that of the

TABLE 1 | The means and standard deviations of the perceived disturbance in the eight different SRT conditions.

	Swedish target				English target			
	Stat	**Fluc**	**BS**	**BE**	**Stat**	**Fluc**	**BS**	**BE**
HI	6.5(1.7)	6.8(1.7)	7.9(1.2)	7.0(1.6)	7.0 (1.7)	7.1 (1.5)	6.9 (1.4)	7.5 (1.5)
NH	6.0(1.8)	6.3(1.8)	6.7(1.8)	6.3(1.7)	6.2(1.8)	6.3(1.9)	7.3(1.8)	6.8(2.1)

HI, Hearing-impaired listeners; NH, Normal-hearing listeners; Stat, Stationary noise; Fluc, Fluctuating noise; BS, Babble Swedish; BE, Babble English.

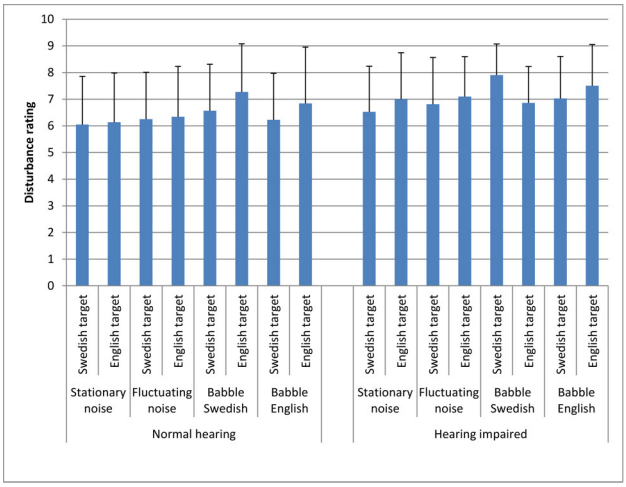


FIGURE 2 | Means and SDs of the perceived disturbance ratings for the NH and HI participants in Swedish and English target with stationary noise, fluctuating noise, babble Swedish, and babble English.

stationary and the *fluctuating* maskers for *English target language* (**Figure 2**).

Discussion

The main aim of this study was to explore how NH and HI listeners perceived disturbance in four different background conditions in their native and non-native languages, respectively. We expected the HI listeners to experience more listening disturbance than the NH listeners. This was not the case, as the current data did not show a statistically significant difference in perceived disturbance between the HI and the NH listeners, although a trend was observed ($p = 0.06$) with relatively high disturbance ratings by the HI listeners.

We also expected the speech maskers to be perceived as more disturbing than the noise maskers. The result confirmed our prediction for the NH listeners but not for the HI listeners. Although the HI listeners perceived a high level of disturbance from the Swedish babble for Swedish as target speech, the Swedish babble for English target speech was not perceived as more disturbing than the other maskers, including the noise maskers. For English as target speech, the NH listeners perceived the Swedish babble as more disturbing than both noise maskers. The characteristics of the maskers applied in the current study seem to affect the perceived disturbance differentially in HI and NH listeners.

Generally, the disturbing effects of interfering speech can be explained in terms of two mechanisms. First, linguistic similarity (Brouwer et al., 2012) between the target speech and the masker speech affect the degree of disturbance, and secondly, the intelligibility of the words in the masker speech affects masker disturbance. Additionally, the disturbing effect of interfering speech is commonly ascribed to higher cognitive processing levels than that of interfering noise. Interfering speech captures attention, induces semantical interference, and is often associated with increased cognitive load (Cooke et al., 2008; Mattys et al., 2009; Koelewijn et al., 2012).The degree of disturbance seems to depend on the lexical familiarity with the masker. Larger interference is observed if the masker has semantically noticeable meaning (c.f., cocktail party effect, Cherry, 1953). The NH listeners in the current study may have overheard more native, familiar words in the Swedish babble masker than the HI listeners. This may have temporarily captured their attention (Conway et al., 2001). For the English target speech/Swedish babble condition, it may have been cognitively more demanding for the NH listeners to focus on the non-native target speech while trying to inhibit speech in their native or most accomplished language.

Surprisingly, for the HI listeners the same condition (i.e., the English target/Swedish babble) was equally disturbing as the disturbance from the other three maskers for English target speech. For the HI listeners, the specific features of the different maskers do not result in differences in perceived disturbance for this non-native target speech: the masking effects of the four maskers are equivalent. One inference to be drawn from this is that the HI listeners most likely had difficulties to perceive any words from the speech maskers correctly. Therefore, the Swedish babble in the English target speech condition was not more disturbing than the other maskers. We also suggest that the HI listeners may have to invest all their processing resources (Rabbitt, 1991) to focus on the English target speech, trying to identify the words and solve the assigned task to listen to and repeat the sentences.

As mentioned earlier, the Swedish target/Swedish babble condition was the most disturbing for the HI listeners. The lack of hearing acuity is likely one reason for this result, as the impaired spectral and temporal resolution (Moore, 1985) causes a reduced ability to distinguish different sounds. Additionally, impaired spectral and temporal resolution increases the difficulty to distinguish the linguistically similar (Brouwer et al., 2012) target and masker speech. The relative similarity between the target and the masker depends on factors like phonological, semantic and/or syntactic content of the two streams. From the English target/Swedish babble condition, we suggest that HI listeners likely did not correctly perceive many words in the masker. Additionally, we suggest that the Swedish target/Swedish

babble condition taps into the same level of phonological and syntactic processing and therefore produces a high level of perceived disturbance for the HI listeners.

Listeners often have better speech perception for relatively unfamiliar maskers as compared to familiar, or intelligible, native speech (e.g., Calandruccio et al., 2010). For the subjectively perceived disturbance ratings, the HI listeners obtained benefit in the Swedish target speech, as the unfamiliar masker (the English babble) was not perceived as more disturbing than the stationary and the fluctuating noise. In the English target speech, the English babble was not perceived as more disturbing than any of the other maskers. The NH listeners did not perceive the English babble as more disturbing than any of the other maskers in the Swedish target speech. However, in English target speech there was no difference between the speech maskers, as the NH listeners perceived both speech maskers (familiar and unfamiliar) as more disturbing than the two noise maskers.

Conclusion

There is no difference in the perceived disturbance from noise and speech maskers during native and non-native speech perception between HI and NH listeners.

For NH listeners, the perceived disturbance from the speech maskers was larger than that from the noise maskers. For HI listeners, the perceived disturbance from speech maskers was similar to that from the noise maskers.

The characteristics of the different masker types applied in the current study seem to influence the perceived disturbance differently in HI as compared to NH listeners.

References

Agus, T. R., Akeroyd, M. A., Gatehouse, S., and Warden, D. (2009). Informational masking in young and elderly listeners for speech masked by simultaneous speech and noise. *J. Acoust. Soc. Am.* 126, 1926–1940. doi: 10.1121/1.3205403

Brouwer, S., Van Engen, K. J., Calandruccio, L., and Bradlow, A. R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners: language familiarity and semantic content. *J. Acoust. Soc. Am.* 131, 1449–1459. doi: 10.1121/1.3675943

Calandruccio, L., Brouwer, S., Van Engen, K. J., Dhar, S., and Bradlow, A. R. (2013). Masking release due to linguistic and phonetic dissimilarity between the target and masker speech. *Am. J. Audiol.* 22, 157–164. doi: 10.1044/1059-0889(2013/12-0072)

Calandruccio, L., Dhar, S., and Bradlow, A. R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. *J. Acoust. Soc. Am.* 128, 860–869. doi: 10.1121/1.3458857

Cherry, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229

Clark, J. G. (1981). Uses and abuses of hearing loss classification. *Asha* 23, 493–500.

Conway, A. R. A., Cowan, N., and Bunting, M. F. (2001). The cocktail party phenomenon revisited; The importance of working memory capacity. *Psych. Bull. Rev.* 8, 331–335. doi: 10.3758/BF03196169

Cooke, M., Garcia Lecumberri, M. L., and Barker, J. P. (2008). The foreign language cocktail party problem: energetic and informational masking effects in non-native speech perception. *J. Acoust. Soc. Am.* 123, 414–427. doi: 10.1121/1.2804952

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., and Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol. Methods* 1, 170–177. doi: 10.1037/1082-989X.1.2.170

Festen, J. M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 88, 1725–1736. doi: 10.1121/1.400247

Gautreau, A., Hoen, M., and Meunier, F. (2013). Let's all speak together! exploring the masking effects of various languages on spoken word identification in multi-linguistic babble. *PLoS ONE* 8:e65668. doi: 10.1371/journal.pone.0065668

Hällgren, M., Larsby, B., and Arlinger, S. (2006). A Swedish version of the hearing in noise test (HINT) for measurement of speech recognition. *Int. J. Audiol.* 45, 227–237. doi: 10.1080/14992020500429583

Hua, H., Emilsson, M., Kähäri, K., Widén, S., Möller, C., and Lyxell, B. (2014). The impact of different background noises: effects on cognitive performance and perceived disturbance in employees with aided hearing impairment and normal hearing. *J. Am. Acad. Audiol.* 25, 859–868. doi: 10.3766/jaaa.25.9.8

Hua, H., Karlsson, J., Widén, S., Möller, C., and Lyxell, B. (2013). Quality of life, effort and disturbance perceived in noise: a comparison between employees with aided hearing impairment and normal hearing. *Int. J. Audiol.* 52, 642–649. doi: 10.3109/14992027.2013.803611

Jahncke, H., and Halin, N. (2012). Performance, fatigue and stress in open-plan offices: the effects of noise and restoration on hearing-impaired and normal-hearing individuals. *Noise Health* 14, 260–272. doi: 10.4103/1463-1741.102966

Kane, M. J., Bleckely, M. K., Conway, A. R. A., and Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *J. Exp. Psychol. Gen.* 130, 169–183. doi: 10.1037/0096-3445.130.2.169

Kilman, L., Zekveld, A. A., Hällgren, M., and Rönnberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Front. Psychol.* 5:651. doi: 10.3389/fpsyg.2014.00651

Kilman, L., Zekveld, A. A., Hällgren, M., and Rönnberg, J. (2015). Native and non-native speech perception by hearing-impaired listeners in noise and speech maskers. *Trends Hear*. 19: 2331216515579127. doi: 10.1177/2331216515579127

Kjellberg, A., Landström, U., Tesarz, M., Söderberg, L., and Åkerlund, E. (1996). The effects of nonphysical noise characteristics, ongoing task and noise sensitivity on annoyance and distraction due to noise at work. *J. Environ. Psychol.* 16, 123–136. doi: 10.1006/jevp.1996.0010

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., and Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. *Hear. Res.* 312, 114–120. doi: 10.1016/j.heares.2014.03.010

Koelewijn, T., Zekveld, A. A., Festen, J. M., and Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear Hear.* 32, 291–300. doi: 10.1097/AUD.0b013e3182310019

Kramer, S., Kapteyn, T., Festen, J., and Kuik, D. (1997). Assessing aspects of auditory handicap by means of pupil dilation. *Int. J. Audiol.* 36, 155–164. doi: 10.3109/00206099709071969

Larsby, B., Hällgren, M., Lyxell, B., and Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *Int. J. Audiol.* 44, 131–143. doi: 10.1080/14992020500057244

Mattys, S. L., Brooks, J., and Cooke, M. (2009). Recognizing speech under a processing load: dissociating energetic from informational factors. *Cogn. Psychol.* 59, 203–243. doi: 10.1016/j.cogpsych.2009.04.001

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse condition: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006

McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., and Wingfield, A. (2005). Hearing loss and perceptual effort: downstream effects on older adults' memory for speech. *Q. J. Exp. Psychol.* 58A, 22–33. doi: 10.1080/02724980443000151

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., et al. (2014). Listening effort and fatigue: what exactly are we measuring? A british society of audiology cognition in hearing special interest group "white paper." *Int. J. Audiol.* 53, 433–445. doi: 10.3109/14992027.2014.890296

Moore, B. C. J. (1985). Frequency selectivity and temporal resolution in normal and hearing-impaired listeners. *Br. J. Audiol.* 19, 189–201.

Moore, B. C., and Glasberg, B. R. (1998). Use of a loudness model for hearing-aid fitting. I. Linear Hearing Aids. *Br. J. Audiol.* 32, 317–335. doi: 10.3109/03005364000000083

Murphy, D. R., Craik, F. I., Li, K. Z., and Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* 15, 323–334. doi: 10.1037/0882-7974.15.2.323

Nachtegaal, J., Kuik, D. J., and Anema, J. R. (2009). Hearing status, need for recovery after work, and psychosocial work characteristics: result from an Internet based National Survey on Hearing. *Int. J. Audiol.* 48, 684–691. doi: 10.1080/14992020902962421

Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. Am.* 95, 1085–1099. doi: 10.1121/1.408469

Pichora-Fuller, K. M. (2009). How cognition might influence hearing aid design, fitting, and outcomes. *Hear. J.* 62, 32. doi: 10.1097/01.HJ.0000364274.44847.dc

Plomp, R., and Mimpen, A. M. (1979).Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18, 43–52. doi: 10.3109/00206097909072618

Rabbitt, P. M. A. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otol. Suppl.* 476, 167–176. doi: 10.3109/00016489109127274

Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (2005). Release from informational masking by time reversal of native and non-native interfering speech. *J. Acoust. Soc. Am.* 118, 1274–1277. doi: 10.1121/1.2000751

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42, 68–76. doi: 10.3109/14992020309074626

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, 99–105. doi: 10.1080/14992020802301167

Skagerstrand, Å., Stenfelt, S., Arlinger, S., and Wikström, J. (2014). Sounds perceived as annoying by hearing-aid users in their daily soundscape. *Int. J. Audiol.* 53, 259–269. doi: 10.3109/14992027.2013.876108

Sörqvist, P., and Rönnberg, J., (2014). Individual difference in distractability: an update and a model. *PsyCh. J.* 3, 42–57. doi: 10.1002/pchj.47

Sörqvist, P., Stenfelt, S., and Rönnberg, J. (2012). Visual-Verbal cognitive load modulates auditory-sensory gating in the brainstem. *J. Cogn. Neurosci.* 24, 2147–2154. doi: 10.1162/jocn_a_00275

Tun, P. A., McCoy, S., and Wingfield, A. (2009). Aging, hearing acuity, and the attentional cost of effortful listening. *Psychol. Aging* 24, 761–766. doi: 10.1037/a0014802

Van Engen, K. J. (2010). Similarity and familiarity: second language sentence recognition in first and second-language multi-talker babble. *Speech Commun.* 52, 943–953. doi: 10.1016/j.specom.2010.05.002

Van Engen, K. J., and Bradlow., and A. R. (2007). Sentence recognition in native and foreign language multi-talker background noise. *J. Acoust. Soc. Am.* 121, 519–526. doi: 10.1121/1.2400666

Venetjoki, N., Kaarlela-Tuomaala, A., Keskinen, E., and Hongisto, V. (2006). The effect of speech and speech intelligibility on task performance. *Ergonomics* 49, 1068–1091. doi: 10.1080/00140130600679142

Versfeld, N. J., and Dreschler, W. A. (2002). The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *J. Acoust. Soc. Am.* 111, 401–408. doi: 10.1121/1.1426376

Zekveld, A. A., Kramer, S., and Festen, J. (2010). Pupil response as an indication of effortful listening. The influence of sentence intelligibility. *Ear Hear.* 31, 480–490. doi: 10.1097/AUD.0b013e3181d4f251

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., van Beek, J. H. M., and Rönnberg, J. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentence in noise. *Ear. Hear.* 32, e16–e25. doi: 10.1097/AUD.0b013e318228036a

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kilman, Zekveld, Hällgren and Rönnberg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 14 September 2015
doi: 10.3389/fpsyg.2015.01390

Speech intelligibility and recall of first and second language words heard at different signal-to-noise ratios

Staffan Hygge, Anders Kjellberg and Anatole Nöstl*

Environmental Psychology, University of Gävle, Gävle, Sweden

Free recall of spoken words in Swedish (native tongue) and English were assessed in two signal-to-noise ratio (SNR) conditions (+3 and +12 dB), with and without half of the heard words being repeated back orally directly after presentation [shadowing, speech intelligibility (SI)]. A total of 24 word lists with 12 words each were presented in English and in Swedish to Swedish speaking college students. Pre-experimental measures of working memory capacity (operation span, OSPAN) were taken. A basic hypothesis was that the recall of the words would be impaired when the encoding of the words required more processing resources, thereby depleting working memory resources. This would be the case when the SNR was low or when the language was English. A low SNR was also expected to impair SI, but we wanted to compare the sizes of the SNR-effects on SI and recall. A low score on working memory capacity was expected to further add to the negative effects of SNR and language on both SI and recall. The results indicated that SNR had strong effects on both SI and recall, but also that the effect size was larger for recall than for SI. Language had a main effect on recall, but not on SI. The shadowing procedure had different effects on recall of the early and late parts of the word lists. Working memory capacity was unimportant for the effect on SI and recall. Thus, recall appear to be a more sensitive indicator than SI for the acoustics of learning, which has implications for building codes and recommendations concerning classrooms and other workplaces, where both hearing and learning is important.

Keywords: noise, recall, speech intelligibility, word lists, signal-to-noise ratio, working memory, working memory capacity

OPEN ACCESS

Edited by:
Carine Signoret, Linnaeus Centre for Hearing and Deafness, Sweden

Reviewed by:
Qinghua He, Southwest University, China
Viveka Lyberg Åhlander, Lund University, Sweden

****Correspondence:***
Staffan Hygge, Environmental Psychology, University of Gävle, SE-801 76 Gävle, Sweden
staffan.hygge@hig.se

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *31 January 2015*
Accepted: *31 August 2015*
Published: *14 September 2015*

Citation:
Hygge S, Kjellberg A and Nöstl A (2015) Speech intelligibility and recall of first and second language words heard at different signal-to-noise ratios. Front. Psychol. 6:1390. doi: 10.3389/fpsyg.2015.01390

Introduction

When the teacher's speech signal is degraded by the acoustic properties of the classroom, speech intelligibility is reduced, which in turn makes learning more difficult. In order to minimize acoustic disturbances in the classroom, government agencies have established building codes, standards, and recommendations for acceptable signal-to-noise ratios (SNRs) and reverberation time in classrooms and other work places, where it is important to hear and understand auditory information (American National Standards Institute, 2002; Vallet and Karabiber, 2002; Swedish Work Environment Authority, 2006, 2011; Swedish Standards Institute, 2007). These codes and standards are based on what is required for correct identification of spoken words or isolated sentences, i.e., speech intelligibility (SI), which mostly is defined as percentage or probability of correct identifications.

However, SI or correct identification of the spoken word is only one factor in memorizing the information and probably not enough. Acceptable listening conditions are no guarantee for good learning. Kjellberg (2004) argued that if the acoustic conditions or other factors make listening harder or requiring more effort, the recall will suffer even if SI is at an acceptable level. The key factor for the impaired recall seems to be that when the limited working memory capacity is depleted, less time and resources are left for processing and storing of the material to be remembered. In two experiments, Kjellberg et al. (2008) and Ljung and Kjellberg (2009) found support for this hypothesis. Similar results have also been reported in other recent papers from our group (Ljung et al., 2009, 2013), as well as by others in earlier studies (Rabbitt, 1966, 1968; Surprenant, 1999, 2007).

One implication of these results is that SI may be a cruder indicator of the quality of the listening conditions than the memory and recall of the spoken message. In order to show that, SI and recall should be assessed independently for the same material and by the same subjects. In earlier studies from our group (Kjellberg et al., 2008; Ljung and Kjellberg, 2009) the participants shadowed all the words they heard in the wordlist. This was done to ensure that the words were captured correctly also in the less favorable listening conditions.

Previous research also indicates that shadowing suppresses the free recall of the early items of on a word list (Petrusic and Jamieson, 1978; cf. also Parkinson et al., 1971; Parkinson, 1972). In our context it was further important to explore whether shadowing influenced the effect on recall also in an unfavorable listening conditioning, such as +3 dB SNR, and whether recall may be an advantage of recall for late items in a wordlist, resulting in over- or underestimation of recall.

A related issue is whether SNR has a more pronounced effect on the memory of second language words than on the native tongue words, even if the SI is equal. This would be expected if encoding of the second language words requires more processing resources, i.e., will be more taxing for working memory.

SNR also interacts with the position of the word in a wordlist. Kjellberg et al. (2008) found that the free recall decreased in the primacy and recency parts of spoken word lists when the words were presented at a lower SNR. As primacy and recency effects are assumed to reflect long- and short-term memory respectively, we wanted to further explore the extent to which SNR had different effects for long- and short-term memory.

The present experiment was designed to investigate these questions. For the recall task four variables were selected as within person factors: (i) whether the spoken wordlists were in Swedish (native tongue) or in English (Language), (ii) whether the words were heard under acceptable or less than acceptable SNR (+12 or +3 dB), (iii) whether the spoken words were shadowed orally directly after presentation or not (Shadowing), and (iv) whether the presented word was in the first, second or third part of the word list (Part). Thus, all participants encountered all experimental combinations of Language, SNR, Shadowing, and Part. In addition, the outcome of a pre-experimental working memory operation span (OSPAN) task was split by the median and included as an between person independent variable of working memory capacity. In a previous study (Ljung et al., 2013), OSPAN was reported to be related to recall, but not to SI.

For the SI, SI task, the probability of correctly identified words in the shadowing task, was analyzed with Language, SNR, and OSPAN as independent variables. In the SI task, the factor Part was not meaningful as the participants repeated back each word immediately after hearing it.

For SI we expected main effects of SNR and Language, but conjectured that the size of the SNR-effect would be higher for recall than for SI.

For the recall of the words the basic hypotheses were that for the +3 dB compared to +12 dB SNR, recall would be worse, which would also be the case for English words compared to the Swedish words. The size of the loss in recall from SNR for English words was expected to be larger than for Swedish words. Our OSPAN measure of working memory capacity was expected to show up both as a main effect and in interactions with SNR or Language.

Materials and Methods

Participants

A sample of 48 undergraduate students with a mean age of 27.1 years (SD = 7.8) and with equal numbers of men and women participated in the study. They were recruited by information screens in the university premises. Self-reported normal hearing, reading and writing skills were inclusion criteria and the subjects received a cinema ticket for their participation. All participants had studied English for 9 or 10 years before they entered university studies, at which level most readings for their courses are in English. Thus, their English proficiency is quite high. None of the participants had taken English at university level. On arrival the participants were informed about the study, and about their right to leave the experiment at any time without giving any reason. On a direct question all the subjects agreed to participate. For this research we have an ethical approval from the Regional Ethical Board in Uppsala (Nr 338/2011), which allows to take an informed verbal consent, rather than a written one, given that it is documented by whom, to whom, where and when the consent was given. This was done in the present study.

Word Lists

Twenty-four word lists with twelve words each were generated, twelve lists in English and 12 in Swedish. The words were taken from 24 semantic categories and chosen from category norms for the two languages in which the words are ranked with respect to the strength of their association with the category. For the Swedish words category norms reported by Nilsson (1973) and Hellerstedt et al. (2012) were consulted. For the English words we relied on works by Battig and Montague (1969), Posnansky (1978), Marshall and Parr (1996) and van Overschelde et al. (2004).

Because the subjects were native Swedish speakers, the English word lists were slightly modified to reduce any SI difference between the Swedish and English lists. A few English words which were judged by the first author to be uncommon to the participants were replaced with more common ones. The average

number of syllables were about the same for the English and Swedish words ($F < 1$, means English: 1.62, Swedish: 1.65) and there was no significant interaction between Language and Part of the word lists in this respect. Thus, it can be said that the difference in difficulty between the English and Swedish words lists were not a mere reflection of the length of the words and the number of syllables.

The average category norm rank orders of the individual words were made equal for all the lists (Graeco-Latin squares). Three sets with eight lists each and with counterbalanced presentation orders in the eight combinations of language, SNR, and shadowing were generated. The words were recorded in one session from a female speaker, fluent in both English and Swedish, in a sound-attenuated chamber and normalized to 66 dB(A). The words were read to the participants with a 3 s interval between the words. Broad band noise was added to the word lists to create the SNR conditions of +12 and +3 dB. The lists were presented to participants via Sennheiser HD-202 headphones. All the equipment, including computers (Dell) that the participants used was of the same make and model.

Participants were instructed to memorize as many of the spoken words as possible. After each list participants were given 1 min to type down the words they could remember from the most recent list. The computer model was the same for all participants. This procedure continued until all 24 lists had been presented. The probability for recall of the presented words was the basis dependent measure, and the participants were given a score of 1 for each correctly recalled word even if the spelling was not perfect.

For the half of the word lists that made up the SI shadowing task, 12 in each language, the participants were instructed to repeat aloud the words they heard (shadowing). The lists where shadowing occurred were counterbalanced for the presentation orders in the crossed combinations of SNR and Language. The participants' verbal responses were tape recorded and they were given 0 or 1 as probability scores for the 12 words in each word list, even if the pronunciation was not perfect.

Operation Span

A Swedish translation of the automated OSPAN task (Unsworth et al., 2005) was administered as a pre-experimental measure of working memory capacity. Mathematical operations (e.g., "Is $(5 + 3) \times 3 = 24$?") were presented on a computer screen. The participant was told to respond "yes" or "no" to the operation, as quickly as possible, by pressing a button on the screen using the computer mouse. When a response was recorded, a letter was presented for 0.8 s and the participant was told to remember it for later recall. After that a new mathematical operation was presented or the list ended. The list lengths varied between 3 and 7 letters. A total of 15 lists were used (3 of each list length), and the length increased across the task. When a list ended, the participants were asked to recall the letters in order of presentation. Points were given for each word recalled in the correct serial position and the score for each list was multiplied by the length of the list in order to balance differences in list difficulty. The accumulated points were divided by the total amount of lists (i.e., 15), yielding a maximum possible score of 27 (the maximum observed score was 26.5).

Procedure

Between one and three participants were tested in each session. Each participant was seated with headsets on in front of an individual laptop in a sound attenuated test-room. All participants started with the self-paced OSPAN task.

After the OSPAN task the participants adjusted the listening level in the headphones to a comfortable level, and began with a training phase in which they listened to two lists each from the two languages, with the two SNR levels crossed with the two levels of shadowing. After the training phase the 24 wordlists were presented. The duration of each word was approximately 1 s with a 3-s interstimulus interval. The presentation order of the lists was pseudo-randomized and counter balanced for each set of eight participants. The window for typing in the recalled words remained open for 60 s and was followed by the playback of the next list. The total session lasted 55–65 min depending on how fast participants completed the OSPAN task.

Statistical Analyses

The OSPAN scores were split by the median to form one group with high OSPAN-scores and one group with low scores (Means: High – 22.05, Low – 15.20; *SD*: High – 2.45, Low – 2.65).

For the analysis of SI-shadowing a split-plot ANOVA was performed with Language and SNR as within-subject factors and OSPAN as a between-subject factor. For the analysis of the recall scores Shadowing and the three Parts of the wordlists were added on as within-person variables. That is, position 1–4 in the list were defined as Part 1, position 5–8 as Part 2, and position 9–12 as Part 3.

Thus, separate ANOVAs were run for SI and recall, not a grand MANOVA for them together because we had SI scores for only half of the lists and also that the variable Part did not make sense in the immediate response asked for in the SI-task.

Results

When reporting the results, decimals in the degrees of freedom for the *F*-tests indicate that a Greenhouse–Geisser correction was made because of violations of the sphericity assumption.

SI-Shadowing

In the SI shadowing task, three participants (two males, one female) were excluded because of recording errors or for not following instructions. There was no main effect of OSPAN on SI [$F(1,43) = 0.482$, $p > 0.10$], and no significant interactions between OSPAN and the other independent variables or their combinations (all ps > 0.10) and, therefore, the subsequent SI-analyses were performed without the OSPAN dichotomization, and with Language and SNR as the independent variables. The was a significant main effect of SNR [$F(1,44) = 11.63$, $p < 0.001$, Cohens $d = 0.50$, Means: +3 dB = 11.13, +12 dB = 11.60] indicating more of the 12 words in each list was correctly shadowed with the higher dB-value. There was no significant

main effect of Language [$F(1,44) = 2.26$, $p > 0.10$], and there was no significant interaction SNR × Language. Thus, for SI there was only a marked main effect of SNR with a medium effect size. The lack of any effects of Language strongly indicates that the Swedish and English lists did not differ in SI.

Free Recall

Also for the free recall task there was no main effect of OSPAN [$F(1,46) = 1.05$, $p > 0.10$; **Table 1**]. An inspection of all the interactions between OSPAN and all the other four independent variables in all 15 combinations only yielded one single significant interaction, Language × Part × OSPAN ($p = 0.046$), which was deemed to be of minor importance and being too close to what 5% pure chance mass significance would yield. Thus, to increase group size, power, reliability and sensitivity the subsequent analyses of recall were made without the OSPAN factor, leaving Language, Shadowing, SNR and Part as the independent variables for the free recall task.

For the free recall task the main effects are shown in **Table 1**. Note the high Cohen *d* for SNR (1.01), which is noticeably higher than for the SI-shadowing task above ($d = 0.50$), and the close to strong effect of Language (0.72).

Table 2 and **Figures 1** and **2** show the resulting significant interactions between our experimental variables on recall.

As seen from the general form of the curves in **Figure 1**, recall is best at the end of the word list (recency effect), and second best at the beginning of the list (primacy effect). This reflects the well-known serial position effect. **Figure 1** also shows the significant interactions SNR × Part, and Shadow × Part, and the numerical details of these interactions are given in **Table 2**

Figure 1A indicates that that the higher SNR makes a positive difference at the beginning and at the end of the word lists, but not in the middle of the wordlist. A test of simple main effects of SNR in the three Parts of the word lists revealed significant effects of SNR in Part 1 ($p < 0.000$) and Part 3 ($p < 0.000$), but not in Part 2 ($p = 0.947$). That is, the higher (+12 dB) SNR value was an advantage in the first and last parts of the lists, but not in the middle part.

Figure 1B shows the shifting advantage from shadowing the words. In the first two parts shadowing *impaired* recall of the words, but in the last part there was an advantage of having repeated the words. A test of simple main effects of Shadowing in the three parts of the word list showed significant effects (all *p*s < 0.005) for all three pairwise comparisons, but the direction of the differences changed in the third part of the list. Thus, shadowing the words interfered with, rather than enhanced the subsequent recall of the words in the first two-thirds of the list.

Figure 2 shows the significant three-way interaction Language × Shadowing × SNR. For the English words lists, there was a significant simple main effect of shadowing at SNR +3 dB ($p < 0.005$), but not at SNR +12 dB ($p = 0.318$). For the Swedish word lists there were no significant simple main effect of shadowing neither at SNR +3 dB ($p < 0.723$), nor at SNR +12 dB ($p = 0.088$). Thus, shadowing seems to be a more important negative variable for the recall of English word lists, than for the native tongue Swedish word lists, in its effect on recall at +3 dB.

In summary, the main findings were that both SI and recall was impaired in the unfavorable listening condition (+3 dB), but the effect size was larger for recall than for SI. Language also had a main effect on recall, with a medium effect size, but Language did not have any significant effect on SI. Further, the effect of shadowing on recall was negative for the first two parts of the list, but positive for the last part. Shadowing had no general effect on the effect of SNR on recall, but for the English word lists it added to the negative effect in the +3 dB condition.

Discussion

A notable feature in the results is the difference between the performance on the SI in the shadowing task and the free recall of the words. For the variables that were the same across the two tasks, SNR had a strong main effect for both SI and recall, but the effect size for the effect on recall was higher (1.01) than for the effect on SI (0.50). For language there was a marked effect on recall with an effect size of 0.72, which approached a strong effect, but language did not have any significant effect on SI. As there was no difference in SI between the Swedish and English wordlists, the effects reported on recall are not a matter of the participants not having heard the English words as good as the Swedish words. An

TABLE 2 | *F*-ratios for the significant interactions of the independent variables on free recall.

Variance source	*F*-test – Greenhouse–Geisser *df*-adjusted for the sources including *Part*	Power	η_p^2
Shadow × part	$F(1.81,84.90) = 18.29$, $p = 0.000$	1.00	0.280
SNR × part	$F(1.99,93.63) = 7.49$, $p = 0.001$	0.94	0.140
Language × shadow × SNR	$F(1,47) = 5.07$, $p = 0.029$	0.60	0.097

TABLE 1 | Means and *F*-ratios for the main effects on recall.

Variance source	Recall, means probability		*F*-test – Greenhouse–Geisser *df*-adjusted for the source *Part*	Cohen *d* for some of the effects
Language	English: 0.423	Swedish: 0.461	$F(1,47) = 24.743$, $p < 0.001$	0.72
Shadowing	No: 0.454	Yes: 0.430	$F(1,47) = 6.788$, $p = 0.012$.	
OSPAN	Low: 0.420	High: 0.464	$F(1,46) = 1.050$, p = n.s.	
Signal-to-noise ratio (SNR)	+3 dB: 0.418	+12 dB: 0.466	$F(1,47) = 49.403$, $p < 0.001$	1.01
Part	1: 0.394 2: 0.342 3: 0.590		$F(1.29,60.58) = 69.048$, $p < 0.001$	

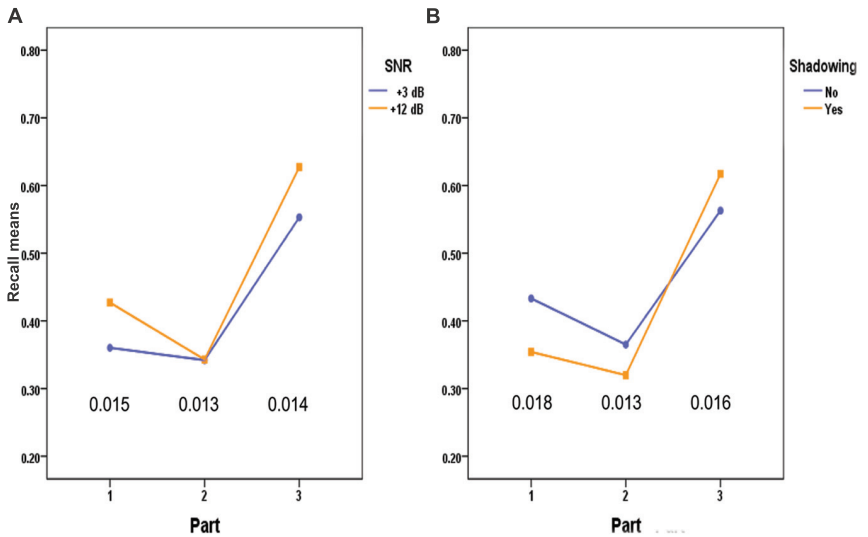


FIGURE 1 | Recall of words in the three Parts of the word lists by SNR (A) and Shadowing (B). The values at the bottom of the figures are the standard errors of the mean differences between the vertically oriented pairs of means.

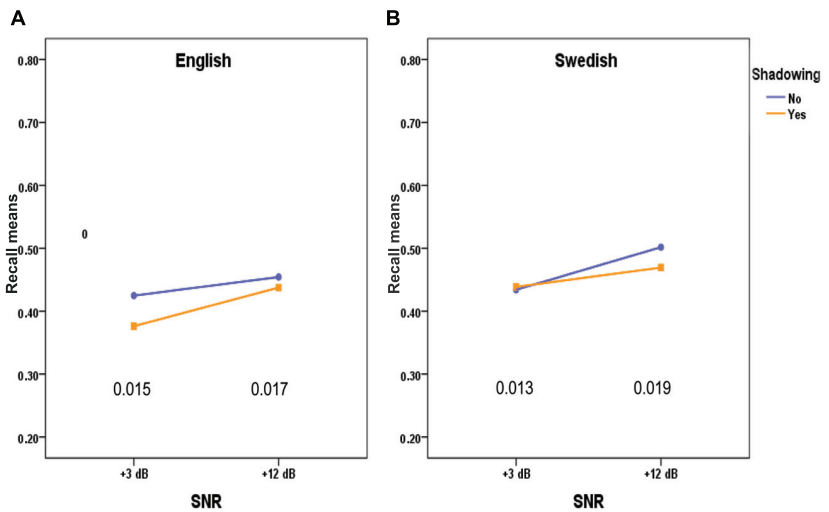


FIGURE 2 | Recall of words in English (A) and Swedish (B) by SNR and Shadowing. The values at the bottom of the figures are the standard errors of the mean differences between the vertically oriented pairs of means.

explanation of the effects on recall then must be sought elsewhere, and our suggestion is centered on the limited capacity of the working memory, which makes it harder to elaborate, analyze and memorize the English words, even if they are as intelligible as the Swedish words.

The results support our basic hypothesis, that recall is a more sensitive indicator than SI when assessing the acceptability of the acoustic conditions in premises, like schools where understanding and memory of spoken information is critical. Thus, it would be more relevant to base acoustic norms and recommendations on memory and recall rather than on SI.

For the recall task, the effects varied between the three parts of the wordlists. The positive effect of the +12 vs. +3 dB SNR was seen both in the first and last part, but not in the middle part. One interpretation of this can be based on what is thought about the nature of the serial recall learning curve, where the early parts of the curve are seen as a consequence of more opportunities for rehearsal, and thereby a more efficient transfer into the long-term memory. The more words that are added to the list, the less are the possibilities to rehearse all preceding words, leading to a less efficient transfer to long term memory. Recall of the last part of the list is assumed to reflect short-term memory. Along this argument it can be argued that the words heard at +3 dB need more working memory resources than the +12 dB words, and thus less capacity is left for storing and retrieval at SNR +3 dB.

A somewhat surprising effect of shadowing was that it had a positive effect on recall only at the end of the lists. The negative effect in the first and second parts is consistent with previous research (Parkinson et al., 1971; Parkinson, 1972; Petrusic and Jamieson, 1978) and seen as an overall negative net effect of shadowing. Shadowing in the first two parts of the wordlists probably impaired recall by interfering with rehearsal of the preceding words. Rehearsal of the words in the last part of the list in memory seem to be less important for recall because they are within the time reach of the echoic memory (there was about 12 s from the first word in Part 3 of the list until typing in the recalled words). Therefore, the elaboration and rehearsal of the words required when shadowing might have had a positive effect on recall.

Shadowing did not have a general effect in the unfavorable listening condition (+3 dB) but the interaction Language × Shadowing × SNR, as depicted in **Figure 2**, suggests that it has such an effect when the list contains second language words. One explanation of this effect is that some of the words in English, which were not more difficult to shadow, still took more of working memory resources than Swedish words at the low SNR-level, which then resulted in inferior recall.

Contrary to the hypothesis, the more unfavorable listening condition did not have a more marked detrimental effect on the memory of the English lists compared to the Swedish ones, indicated by the lack of an interaction Language × SNR. A possible explanation is that the English words were so well-known to the student participants that they were as easily identified as the Swedish ones. The recall of the last words (position 12) in each wordlist under shadowing and at +12 dB SNR did not reveal any significant difference in recall between English and Swedish words (Means 0.85 and 0.88, $F < 1$, and this non-significant difference was true for all the three blocks of list presentations, all pairwise Fs < 1.63). Thus, with the lists used in this study the two languages might have been at approximately the same comprehension level.

From an applied perspective it would have been an advantage to have an estimate of how difficult the English word list were in comparison with the Swedish lists for the group we studied. However, from a basic experimental point of view and in the analysis of variance it is quite admissible to compare levels of independent variables, such as difficulty of English and Swedish words, even if we do not have a magnitude measure of the degree of difference between the two levels.

It can also be argued that the category norms for the English words in our words lists should have been assessed in sample similar to the one we used to avoid the problem that the "true" category norm count for the English words when presented to our Swedish college students may not be the same as for first language English speakers. However, we decided not to do that, because that would a too large project of its own, but in a way we came fairly close to having comparable probabilities between the English and Swedish words as there was no significant effect of Language for the SI-shadowing task. (See Results – SI-shadowing).

From the ecological relevance point of view, the learning of word lists is a rare task outside the laboratory. However, similar effects to those reported here have been shown for memory of lectures listened to in different acoustic conditions (Ljung et al., 2009). A better and ecologically more valid test of the effect of language would probably be to study memory for a text in English and Swedish. In such a situation, it is likely that the interpretation of the meaning of the English text would require more working memory resources, and the difference in recall between the two languages would be more pronounced.

Further studies are wanted to use these results for more direct acoustic recommendations for learning. As of now we can only conclude that recall and memory seem to be a better and more sensitive indicator than SI of the acoustic conditions. However, we do not know the exact range of the SNR to produce decrements in recall. It may well be the case that also a SNR of +12 dB is not the best SNR for good recall.

In a forthcoming study we will have more to say about acoustic conditions and recall of word lists, and whether the introduction of two levels of reverberation times interacts, or not, with the same SNR-levels as used in the present study. Doing that will give more empirical facts in the process of re-evaluating building codes and recommendation for the acoustic conditions in rooms, such as class rooms, where not only listening, but also memory and learning are important.

Acknowledgments

We are grateful to Anders Hurtig and Marijke Keus van de Poll who made the data collection and to Johan Odelius for editing the sound files. The Swedish Research Council Formas contributed to the funding of this study.

References

American National Standards Institute. (2002). *Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools: ANSI S12.60*. Melville, NY: Standards Secretariat Acoustical Society of America.

Battig, W. F., and Montague, W. F. (1969). Category norms for verbal items in 56 categories: a replication and extension of the Connecticut category norms. *J. Exp. Psychol. Monogr.* 80, 1–46. doi: 10.1037/h0027577

Hellerstedt, R., Rasmussen, A., and Johansson, M. (2012). Swedish category norms. *Lund Psychol. Rep*. 12, 1–96.

Kjellberg, A. (2004). Effects of reverberation time on the cognitive load in speech communication: theoretical considerations. *Noise Health* 7, 11–22.

Kjellberg, A., Ljung, R., and Hallman, D. (2008). Recall of words heard in noise. *Appl. Cogn. Psychol.* 22, 1088–1098. doi: 10.1002/acp.1422

Ljung, R., and Kjellberg, A. (2009). Long reverberation time decreases recall of spoken information. *Build. Acoust.* 16, 301–312. doi: 10.1260/135101009790291273

Ljung, R., Israelsson, K., and Hygge, S. (2013). Speech intelligibility and recall of spoken material heard at different signal-to-noise ratios and the role played by working memory capacity. *Appl. Cogn. Psychol.* 27, 198–203. doi: 10.1002/acp.2896

Ljung, R., Sörqvist, P., Kjellberg, A., and Green, A.-M. (2009). Poor listening conditions impair memory for intelligible lectures: implications for acoustic classroom standards. *Build. Acoust.* 16, 257–265. doi: 10.1260/135101009789877031

Marshall, C. E., and Parr, W. V. (1996). New Zealand norms for a subset of Battig and Montague's (1969) categories. *N. Z. J. Psychol.* 25, 24–29.

Nilsson, L. G. (1973). *Category Norms for Verbal Material*. Report 135. Uppsala: Department of Psychology, Uppsala University.

Parkinson, S. R. (1972). Short-term memory while shadowing: multiple-item recall of visually and of aurally presented letters. *J. Exp. Psychol.* 92, 256–265. doi: 10.1037/h0032077

Parkinson, S. R., Parks, T. E., and Kroll, N. E. (1971). Visual and auditory short-term memory: the effects of phonemically similar auditory shadow material during the retention interval. *J. Exp. Psychol.* 87, 274–280. doi: 10.1037/h0030551

Posnansky, C. J. (1978). Category norms for verbal items in 25 categories for children in Grades 2–6. *Behav. Res. Methods Instrum.* 10, 819–832. doi: 10.3758/BF03205407

Petrusic, W. M., and Jamieson, D. G. (1978). Differential interpolation effects in free recall. *J. Exp. Psychol.* 4, 101–109. doi: 10.1037/0278-7393.4.1.101

Rabbitt, P. (1966). Recognition memory for words correctly heard in noise. *Psychon. Sci.* 6, 383–384. doi: 10.3758/BF03330948

Rabbitt, P. (1968). Channel-capacity, intelligibility and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158

Surprenant, A. M. (1999). The effect of noise on memory for spoken syllables. *Int. J. Psychol.* 34, 328–333. doi: 10.1080/002075999399648

Surprenant, A. M. (2007). Effects of noise on identification and serial recall of nonsense syllables in older and younger adults. *Aging Neuropsychol. Cogn.* 14, 126–143. doi: 10.1080/13825580701217710

Swedish Standards Institute. (2007). *Ss 25268:2007 Acoustics - Sound Classification of Spaces in Buildings - Institutional Premises, Rooms for Education, Preschools and Leisure-Time Centers, Rooms for Office Work and Hotels*. Stockholm: Swedish Standards Institute.

Swedish Work Environment Authority. (2006). *Rapport – Verksgemensamt Projekt – Buller I skolmiljö 2003. [Report – Interdepartmental project – Noise in the school environment 2003]*. CTO 2003/482. Stockholm: Swedish Work Environment Authority.

Swedish Work Environment Authority. (2011). Södersten, M., and Lindhe, C: *Kunskapsöversikt - Yrkesrelaterade röststörningar och Röstergonomi*. Stockholm: Arbetsmiljöverket. Report 2011, 6. [Review – Work related voice problems and voice ergonomics].

Unsworth, N., Heitz, R. P., Schrock, J. C., and Engle, R. W. (2005). An automated version of the operation span task. *Behav. Res. Method* 37, 498–505. doi: 10.3758/BF03192720

van Overschelde, J. P., Rawson, K. A., and Dunlosky, J. (2004). Category norms: an update and expanded version of the Battig and Montague (1969) norms. *J. Memory Lang.* 50, 289–335. doi: 10.1016/j.jml.2003.10.003

Vallet, M., and Karabiber, Z. (2002). Some European policies regarding acoustical comfort in educational buildings. *Noise Control Eng. J.* 50, 58–62. doi: 10.3397/1.2839678

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Hygge, Kjellberg and Nöstl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 25 August 2015
doi: 10.3389/fpsyg.2015.01153

The contribution of phonological knowledge, memory, and language background to reading comprehension in deaf populations

Elizabeth A. Hirshorn[1,2], Matthew W. G. Dye[3], Peter Hauser[4], Ted R. Supalla[5] and Daphne Bavelier[1,6]*

[1] Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA, [2] Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, USA, [3] Department of Liberal Studies, National Technical Institute for the Deaf, Rochester Institute of Technology, Rochester, NY, USA, [4] Department of American Sign Language and Interpreting Education, National Technical Institute for the Deaf, Rochester Institute of Technology, Rochester, NY, USA, [5] Department of Neurology, Georgetown University, Washington, DC, USA, [6] Faculté de Psychologie et des Sciences de l'Éducation, Université de Genève, Geneva, Switzerland

OPEN ACCESS

Edited by:
Mary Rudner, Linköping University, Sweden

Reviewed by:
Chloe Marshall, University College London, UK
Jacqueline Leybaert, Université Libre de Bruxelles, Belgium

***Correspondence:**
Elizabeth A. Hirshorn, Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260, USA
hirshorn@pitt.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *27 February 2015*
Accepted: *24 July 2015*
Published: *25 August 2015*

Citation:
Hirshorn EA, Dye MWG, Hauser P, Supalla TR and Bavelier D (2015) The contribution of phonological knowledge, memory, and language background to reading comprehension in deaf populations. Front. Psychol. 6:1153. doi: 10.3389/fpsyg.2015.01153

While reading is challenging for many deaf individuals, some become proficient readers. Little is known about the component processes that support reading comprehension in these individuals. Speech-based phonological knowledge is one of the strongest predictors of reading comprehension in hearing individuals, yet its role in deaf readers is controversial. This could reflect the highly varied language backgrounds among deaf readers as well as the difficulty of disentangling the relative contribution of phonological versus orthographic knowledge of spoken language, in our case 'English,' in this population. Here we assessed the impact of language experience on reading comprehension in deaf readers by recruiting oral deaf individuals, who use spoken English as their primary mode of communication, and deaf native signers of American Sign Language. First, to address the contribution of spoken English phonological knowledge in deaf readers, we present novel tasks that evaluate phonological versus orthographic knowledge. Second, the impact of this knowledge, as well as memory measures that rely differentially on phonological (serial recall) and semantic (free recall) processing, on reading comprehension was evaluated. The best predictor of reading comprehension differed as a function of language experience, with free recall being a better predictor in deaf native signers than in oral deaf. In contrast, the measures of English phonological knowledge, independent of orthographic knowledge, best predicted reading comprehension in oral deaf individuals. These results suggest successful reading strategies differ across deaf readers as a function of their language experience, and highlight a possible alternative route to literacy in deaf native signers.

Highlights:

1. Deaf individuals vary in their orthographic and phonological knowledge of English as a function of their language experience.
2. Reading comprehension was best predicted by different factors in oral deaf and deaf native signers.
3. Free recall memory (primacy effect) better predicted reading comprehension in deaf native signers as compared to oral deaf or hearing individuals.
4. Language experience should be taken into account when considering cognitive processes that mediate reading in deaf individuals.

Keywords: deafness, reading, sign language, orally-trained, short-term memory, phonological awareness, semantic-based memory

Introduction

Learning to read, although a rite of passage for most children, remains a significant educational challenge. It is widely known that learning to read is especially difficult for deaf individuals, with the average deaf reader reaching only a fourth grade reading level (Traxler, 2000). For hearing individuals, foundational steps to achieving skilled reading comprehension include becoming aware that words are made of smaller units of speech sounds, a process termed phonological awareness, and then learning to link visual and phonological information to decode print into already known spoken words (Wagner and Torgesen, 1987; Stahl and Murray, 1994; Høien et al., 1995). As they are sounded out, words are then mapped onto their existing semantic representations and knowledge of the syntax and regularities of the language then help the extraction of meaning from text (Wagner and Torgesen, 1987; Cornwall, 1992; Wagner et al., 1994; Hogan et al., 2005). In deaf populations, where there is not necessarily a known spoken language to map the print information onto, becoming a proficient reader poses its own set of challenges. In this study, we ask which component processes mediate reading comprehension in deaf individuals with severe-to-profound hearing loss, and in particular, investigate the impact of phonological knowledge, memory processes and language experience on reading comprehension (Fletcher, 1986; Wagner and Torgesen, 1987; Swanson, 1999; Swanson and Ashbaker, 2000; Scarborough, 2009).

A main determinant of reading in hearing populations remains the mastery of phonological awareness skills, especially those measured at the single word level (Wagner and Torgesen, 1987; Hatcher et al., 1994; Wagner et al., 1994). In young readers, strong phonological representations facilitate word identification skills, which support comprehension (Perfetti and Hart, 2001; Perfetti et al., 2005). Thus, phonological awareness often comes to predict text comprehension (Shankweiler and Liberman, 1989; Hatcher et al., 1994; Wagner et al., 1994), although the role of phonological awareness in reading skill generally decreases with age (Wagner et al., 1997; Parrila et al., 2004). Nevertheless, phonological coding during comprehension can persist into adulthood (Coltheart et al., 1988) and also continues to be linked to reading skill in reading disorders (Bruck, 1992; Elbro et al., 1994; although see Landi, 2010). Accordingly, phonological deficits are often at the source of reading problems (Pennington and Bishop, 2009) and believed to be a main predictor of reading deficits like dyslexia (Snowling, 1998; Gabrieli, 2009). Phonological remediation, or explicit phonological awareness training, often helps to improve reading skill in dyslexic readers, at least when measured at the word level (Eden et al., 2004; Shaywitz et al., 2004).

Despite clear reasons why the link between English phonological knowledge and reading comprehension may be different in deaf individuals with impoverished access to auditory signals, the main focus in most research on reading in the deaf has been based on the established hearing model of reading, which emphasizes the role of phonological processing. However, it is still unclear whether phonological awareness of English is similar in deaf and hearing individuals or used in the same way to facilitate reading (Mayberry et al., 2011; Bélanger et al., 2012a), depending on how it is acquired (LaSasso et al., 2003). An inherent complication is that most standard tasks used to evaluate phonological knowledge in hearing populations require speech production; yet, many deaf individuals are not at ease with vocalizing English. Based on the many strategies for completing a speech-based phonological assessment used in the literature, it remains unclear whether deaf individuals have qualitatively similar phonological awareness of English to that of hearing individuals. It is important to note that deaf individuals have access to other types of phonological knowledge through the use of signed languages. These also have a phonological structure (MacSweeney et al., 2008) that can support higher cognitive processes (Aparicio et al., 2007; MacSweeney et al., 2009; Morford et al., 2011). Given our present focus on what is termed 'phonological awareness' in the reading literature, the term 'phonological' will refer to phonology of spoken English hereafter. We briefly review below the role of English phonological knowledge, memory processes, and language experience on reading in the deaf.

Several groups have found similarities between deaf and hearing participants in English phonological tasks. Hanson and Fowler (1987) examined deaf signers and found that phonological similarity between English word pairs reduced the reading rate in a speeded lexical decision for both the hearing and the signing deaf individuals, concluding that deaf and hearing participants were using a similar phonetic coding strategy. In another study,

Hanson and McGarr (1989) found that signing deaf college students were able to perform a rhyme generation task, but not with the same degree of success as their hearing peers. Sterne and Goswami (2000) argued that deaf readers possess phonological awareness at different levels (i.e., syllable, rhyme, phoneme), although they lagged behind their hearing peers. Nevertheless, a recent meta-analysis by Mayberry et al. (2011) found just as many studies reporting that deaf individuals have phonological awareness as studies that found that they do not.

Large variation in the type of tasks used to assess phonological awareness in the deaf may in part account for this discrepancy (e.g., syllable, phoneme, rhyme; Hanson and Fowler, 1987; Sterne and Goswami, 2000). In addition, some studies have used spoken responses, a standard method used in hearing populations to study phonological awareness (e.g., Luetke-Stahlman and Nielsen, 2003); however, spoken response is potentially problematic, especially for deaf individuals that are not comfortable with vocalizing. Other studies require the manipulation of written words to assess phonological awareness, but doing so inherently involves reading and orthographic processing. To reduce such potential confounds, several studies have adopted picture stimuli and asked for phonological judgments about the English names corresponding to the pictures, which has allowed for a less contaminated measure of English phonological awareness in deaf individuals (Sterne and Goswami, 2000; Dyer et al., 2003; MacSweeney et al., 2008; McQuarrie and Parrila, 2009). These studies suggest some level of phonological awareness in deaf individuals, with some pointing to the importance of orthographic-to-phonological regularities in supporting such knowledge. An important feature of English is that it is an opaque writing system without one-to-one mapping of graphemes to phonemes. There are, however, interesting consistencies in the visual orthography that could lead to alternative visual or orthographic strategies when performing a phonological task (McQuarrie and Parrila, 2009). The extent to which English phonological knowledge in deaf populations is based on orthographic regularities will be examined in Experiment 1. We present novel picture-based tasks, designed to assess English phonological knowledge, with the feature that the orthographic-to-phonological regularity of the test items is systematically manipulated in order to separately assess shallow knowledge (based on orthography) versus deep knowledge (phonological knowledge above and beyond orthography).

While the emphasis on phonological awareness has been productive in motivating best practices in general reading instruction for hearing individuals (Trezek et al., 2010), it may obscure the fact that *comprehension* is the end goal of reading (McCardle et al., 2001). Text comprehension also calls upon more general cognitive processes. Verbal short-term memory has been shown to correlate with reading skill in a wide range of studies (Siegel and Linder, 1984; McDougall et al., 1994; Swanson and Howell, 2001). Serial recall is often used as an assessment of verbal STM, and is known to rely heavily on phonological processes, as exemplified by a rich literature on the phonological loop and its rehearsal mechanism in speakers (Baddeley et al., 1984; Burgess and Hitch, 1999; Melby-Lervåg and Hulme, 2010; Bayliss et al., 2015). Importantly, serial recall and other verbal STM measures have been shown to contribute unique variance in explaining reading skill compared to phonological measures alone, at least in hearing readers (Gathercole et al., 1991; McDougall et al., 1994). A few studies have directly compared short-term memory capacity in deaf and hearing individuals. Studies of either orally trained deaf individuals or deaf native signers suggest a reduced STM span in the deaf, whether tested in English or in American Sign Language (ASL; Conrad, 1972; Bellugi et al., 1975; Boutla et al., 2004; Koo et al., 2008). Evidence suggests that this difference is attributable to language modality rather than sensory deprivation, *per se*, as hearing bilinguals have lower STM span in ASL as compared to when tested in English. The precise source of such span differences remains debated with current hypotheses focusing on lesser reliance on the temporal chunking of units in the visual modality (Hall and Bavelier, 2010; Hirshorn et al., 2012) and on factors that would differentially affect articulatory rehearsal, such as 'heavier' phonological units (Geraci et al., 2008; Gozzi et al., 2011) or more "degrees of freedom" in phonological composition in sign languages (Marshall et al., 2011). Despite the evidence for serial span group differences, working memory capacity, which is vital when reading tasks are more demanding, has been shown to be equal for deaf and hearing individuals (Boutla et al., 2002, 2004).

Free recall memory span has also been linked with overall reading skill and comprehension (Dallago and Moely, 1980; Lee, 1986). In contrast to serial recall, free recall is thought to rely more heavily on semantic processing, with greater time on each item allowing for deeper processing (Craik and Lockhart, 1972; Craik and Tulving, 1975; Melby-Lervåg and Hulme, 2010). Accordingly, performance on free recall tests is improved by semantic relatedness (e.g., Hyde and Jenkins, 1973; Bellezza et al., 1976). Furthermore, in contrast to serial recall that heavily relies on rehearsal mechanisms, free recall tasks have longer post-stimulus delays, which are thought to allow for short-term consolidation that aids memory retrieval (Jolicœur and Dell'Acqua, 1998; Bayliss et al., 2015) although this distinction between serial and free recall continues to be debated (Bhatarah et al., 2009). Free recall also has the added benefit of distinguishing between the primacy (recall of initial list items) and recency effects (recall of last list items), such that primacy effects depend to a larger extent on semantic processing, while recency effects reflect a greater contribution of short-term rehearsal and phonological processing similar to what is observed in serial recall tasks (Martin and Saffran, 1997; Martin and Gupta, 2004). This distinction appears relevant when considering predictors of reading. For example, reading-disabled children have been reported to have a decreased primacy effect, but equivalent recency effect, compared to non-disabled readers (Bauer and Emhert, 1984).

Finally, members of deaf communities typically vary greatly in terms of their language background. While around 48% of deaf or hard-of-hearing children use "speech only" as their main mode of communication (Gallaudet Research Institute, 2005), linguistic knowledge within these individuals varies widely. In addition, many early studies examining reading in deaf individuals did not identify whether deaf participants were native users of a signed language, orally trained or users of other

forms of communication such as Cued Speech or Signed English. This is likely to be important as having access to a *natural* language from birth has been shown to be a precursor to good reading skill in the deaf (Chamberlain and Mayberry, 2000, 2008; Padden and Ramsey, 2000; Goldin Meadow and Mayberry, 2001). Early exposure to a natural language, be it spoken or signed, is associated with better knowledge of grammar and syntax (Mayberry, 1993), executive functioning (Figueras et al., 2008; Hauser et al., 2008a), and meta-linguistic awareness (Prinz and Strong, 1998); all of these in turn appear to foster better reading comprehension (Chamberlain and Mayberry, 2000; Padden and Ramsey, 2000; Goldin Meadow and Mayberry, 2001). For these reasons, we focus here on two distinct groups of deaf readers with early exposure to a natural language: deaf native signers of ASL, who have very limited spoken English skill, and orally trained deaf, that speak and lip-read English and were exposed to speech-based natural language and educated in mainstream schools with hearing peers, termed hereafter *oral deaf*. In Experiment 2, we seek to determine the relative contribution of English phonological knowledge, English orthographic knowledge, serial recall and free recall to reading comprehension in these two populations of deaf readers.

It should be noted that some additional factors naturally co-vary when sampling from these populations. First, despite our selection of individuals with similar *unaided* levels of hearing loss across these two groups, oral deaf individuals are more likely to use hearing aids or have a cochlear implant (CI), which would increase their *aided* hearing loss and access to auditory information. Second, because deaf native signers use ASL as their primary mode of communication, they are more likely to be (bimodal) bilinguals, and also be reading their second language when faced with English text (Chamberlain and Mayberry, 2008; Morford et al., 2011; Piñar et al., 2011). Recent work on reading in deaf native signers suggest, while they clearly possess knowledge of the phonology of English, they may not make use of that phonological knowledge in the same way as hearing individuals do when reading text for comprehension (Miller and Clark, 2011; Bélanger et al., 2012a,b, 2013). It should also be acknowledged that the relative contribution to the reading process of different language experience (such as use of a signed language) and of reading a first versus a second language remains understudied.

In sum, Experiment 1 presents newly developed 'deaf-friendly' measures of English phonology that manipulate whether a 'phonological' task can be solved with an orthographic strategy or not. In doing so, it allows us to separately assess orthographically based phonological knowledge from non-transparent, deep phonological knowledge of English in deaf readers. Experiment 2 then turns to the determinants of reading in our two groups of deaf adults with different language backgrounds by considering the relative contribution of various types of English phonological knowledge that are based upon the phoneme level (both shallow and deep) and larger phonological units (syllable and speechreading measures), linguistic short-term memory (serial recall span) and semantic-based memory (free recall span). Together, this battery is designed to distinguish between various levels of English phonological knowledge and more general cognitive measures as predictors of reading comprehension in our two groups of deaf adults. Based on the existing literature, we predicted weaker deep phonological knowledge in deaf native signers than in the oral deaf. Moreover, we hypothesized that reading comprehension may show a greater reliance on memory processes, especially semantic-based, in deaf native signers, whereas deep phonological knowledge would be the primary predictor of reading skills in the oral deaf.

Experiment 1

The goal of Experiment 1 was to determine the extent and type of English phonological knowledge in two groups of deaf readers. More specifically, we tested the extent to which the two deaf groups utilized visual orthographic knowledge to complete phonological tasks. Two new tests of English phonological knowledge were designed for use with our profoundly deaf participants. An important design feature that was we did not want to require vocal responses or use text-based materials to measure phonological knowledge, making commonly employed tasks like non-word naming inappropriate. Instead our tests require button-press responses and use nameable black and white pictures to provide a cleaner measure of phonological knowledge – there is no explicit phonological representation in the picture itself, unlike for written words. Critically, the transparency of the orthographic-to-phonological mapping was systematically manipulated in order to assess how much a purely orthographic strategy was being used to perform a phonological task. More specifically, the transparency of orthographic-to-phonological mapping was explicitly manipulated such that orthographic information, if used, could either help task performance (shallow task) or be uninformative or counter-productive (deep task). This manipulation was deployed in two separate tasks. The first task required participants to indicate which of three items sounded different from the other two, with the difference being sound-based and located either at the first consonant or vowel. The second task mirrored a phonemic manipulation task often used in the reading literature. Participants were asked to extract the first sound and the last sound of the names corresponding to two pictures, and then combine those to make a new name. We expected to see differences between the deaf groups in the extent to which they utilized an orthographic strategy, with deaf native signers using those strategies more than the oral deaf. We note that a group of hearing participants was also evaluated on these tasks to verify that our stimuli properly assess orthographic and phonological knowledge. Their data are reported in the supplementary information and confirm a gradient from shallow to deep phonology with our materials.

Methods

Participants

The study included 26 profoundly deaf native signers of American Sign Language [M_{age} = 22 (18–32); 17 female; $M_{unaided\ PTA\ loss\ in\ better\ ear}$ = 94 dB, 73–110 dB; Note PTA means Pure Tone Average] and 21 oral deaf (M_{age} = 21 (18–24);

16 female; $M_{unaided\ PTA\ loss\ in\ better\ ear}$ = 90 dB, 63–120 dB). All participants were recruited from the Rochester Institute of Technology (RIT) or the National Technical Institute for the Deaf (NTID).

Inclusion criteria for all participants were: (i) unaided hearing loss of 75 dB or greater in the better ear[1], (ii) onset of deafness before 2 years of age[2], and (iii) being right handed. We were unable to acquire the unaided dB loss level for four oral deaf participants and five of the deaf native signing participants. Based upon deaf participants for whom audiological data was available, the two deaf groups had equivalent levels of unaided dB loss (see **Table 1**). Hearing loss levels were obtained from self-reports as well as consented and IRB-approved access to RIT/NTID records. All participants were treated in accordance with the University of Rochester's Research Subjects Review Board guidelines and were paid for their participation in the study. No participants reported having any learning disorder.

Additional inclusion criteria for deaf native signers included: being born to deaf parents and exposed to ASL from infancy; and having limited spoken English skill, as measured by the TOAL-2 (see below). All deaf native signers reported having used hearing aids at some point in their lives, but only six continued to use hearing aids regularly and three reported using them only occasionally. Twenty of the deaf native signers attended a school for the deaf during at least one phase of their education before college, and six attended a mainstream school throughout.

In contrast, additional inclusion criteria for oral deaf subjects included: being born to hearing parents; being educated in mainstream schools that adopted oral-aural approaches promoting spoken language ability; minimal or absent ASL skills with no exposure to ASL until college years (average of 2.5 years in college; range = 0.5–6 years); using oral communication as the primary mode of communication; and relying on lip-reading to comprehend spoken English. Most of these students received individual speech therapy on a regular basis upon entering the school system and continued to receive speech training and gained skill in speechreading as a part of all of their academic courses. Four of the oral deaf participants had received CI with an age of implantation of 2.5, 5, 17, and 19 years. Of the 17 oral deaf participants without CIs, all wore hearing aids except two. If participants wore CIs or hearing aids, they were instructed to use them as they normally would during all tasks. Six attended a preschool for deaf children, but all attended mainstream schools during their elementary, middle, and high school years. Fourteen participants reported not using ASL at all, while seven reported having some ASL experience starting in college.

In order to verify participants' native language proficiency and to confirm that the groups had distinct and separable language skills, we administered ASL and spoken English proficiency tests that probed both comprehension and production. The American Sign Language Sentence Reproduction Test (ASL-SRT) was used as a test of ASL proficiency (Hauser et al., 2008b; Supalla et al., 2014), and the Test of Adolescent Language Speaking Grammar Subtest (TOAL-2; Hammill, 1987) was used as a test of English proficiency. In both tests, subjects saw/heard sentences of increasing complexity and length and were instructed to repeat back exactly what they saw/heard. Thus, both tests involved both a comprehension and a production component. Only sentences recalled verbatim were counted as correct. Deaf native ASL signers scored the ASL proficiency test (for native signers and oral deaf subjects) and hearing native English speakers scored the English test for oral deaf subjects. The percent accuracy (number of sentences repeated verbatim divided by the total) on each proficiency test was compared between groups (see **Table 1** for mean values). For the spoken English proficiency test, deaf native signers were instructed to respond in ASL if they were not comfortable producing overt speech. Nevertheless, native signers were at floor and therefore a statistical test was not needed. **Table 1** shows performance of the two deaf groups on these two sentence repetition tests. For the ASL-SRT, the native signers were more accurate than oral deaf participants. Overall, the language proficiency results confirmed successful enrollment of two groups of deaf participants with distinct language backgrounds: one group is significantly more skilled in spoken English, and the other more skilled in ASL.

Finally, participants completed the TONI-3 (Brown, 2003) to confirm that the two groups did not have significantly different levels of non-verbal IQ in order to control for the impact of general cognitive factors in reading comprehension. Participants viewed arrays of visual patterns of increasing complexity, with one missing component in each array. They were required to identify the missing component by selecting from 4 or 6 options. Due to a communication error early during data collection, some participants were not given the TONI-3 and thus data are missing for one oral deaf, and six deaf native signers. As can be seen in **Table 1**, TONI-3 scores across groups were not significantly different.

[1]One deaf native signer had an unaided hearing loss of 70 dB and one oral deaf had an unaided hearing loss of 63 dB.

[2]Two oral deaf became deaf at age 4 years.

TABLE 1 | Demographic and language backgrounds of participants (mean scores with ranges or SD).

Measure	Oral Deaf	Deaf native signers	*t*	*df*	*p*
Age	21 (18–24)	22 (18–32)			
dB loss	87 (16)	90 (10)	0.74	36	0.47
TONI standardized score	98.3 (10.6)	99.4 (10.6)	0.34	38	0.73
Native language fluency					
English	31% (18%)	NA			
ASL	5% (9%)	63% (14%)	14.48	45	<0.001

Design and Procedure

The tasks required phonological judgments to be made on the basis of black and white drawings of objects. It was therefore important to ensure that participants knew the desired English names to be associated with the pictures we used. All participants initially named the pictures by typing their corresponding

English name into the computer. There was feedback to make sure they had assigned the correct name and spelling. If a picture was misnamed or misspelled, participants were informed of the mistake and it was presented again at a later time until all pictures had been named and spelled correctly. Instructions were written for oral deaf (and hearing, see Supplemental Information) participants, but the experimenter always reviewed the instructions verbally before the experiment started. An instructional video in ASL was made for signers by a bilingual hearing signer, and gave many examples to ensure the tasks were clear. An ASL/English interpreter skilled in communicating with deaf individuals of varied language background was always present in case clarifications were needed.

Phoneme Judgment Task

The Phoneme Judgment Task employed an 'odd-man-out' paradigm: three pictures were displayed in a triangle formation on a computer screen, and participants were instructed to select the item with a different sound. Participants responded by pressing 'H', 'B', or 'N' on a QWERTY keyword, corresponding to the 'odd-man-out' location on the screen. The odd-man-out could be located either at the first consonant or at the vowel. These two phoneme-type conditions were run blocked with the order of blocks counterbalanced across groups. Words in the first consonant condition could be either one or two-syllables, while the words in the vowel condition were all one-syllable.

The complex letter-to-sound mappings of English were exploited in order to determine whether participants were able to go beyond purely orthographic strategies in order to perform accurately. Two conditions were labeled as "shallow" and these were conditions in which a purely orthographic strategy could yield 100% accuracy. In shallow condition A, the similar sounding pair shared the same orthography whereas the odd-man-out had a different orthography (e.g., **b**elt/dog/door for the first-sound task; k**i**ng/goat/soap for the vowel task). In shallow condition B, 100% accuracy using an orthographic strategy would depend upon flexible letter-to-sound knowledge, such as being aware that 'k' and 'c' can both be mapped to the same sound in English (e.g., **l**emon/kettle/compass for the first-sound task; sk**u**nk/mouse/clown for the vowel task). Another two conditions were labeled as "deep" and were constructed such that accuracy would be poor if an orthographic strategy were employed. In deep condition C, all of the words shared the same letter (e.g., **ch**ef/church/chair for the first-sound task; d**o**ve/rose/cone for the vowel task). This condition therefore requires knowledge of idiosyncratic mappings in English: knowing that 'c' can sometimes sound the same as 's' no longer provides a cue to the correct answer. Finally, deep condition D was constructed such that an orthographic strategy would routinely lead to the incorrect answer. In this condition, the odd-man-out shared orthography with one of the two similar-sounding items (e.g., **k**ey/knee/nurse for the first-sound task; l**ea**f/steak/chain for the vowel task). Examples and more details are provided in **Figure 1**. Before each task, instructions were given using two sample trials. The sample trials contained one 'shallow' and one 'deep' trial to clarify the instructions, but also to demonstrate how they could not always be solved based on orthography alone.

Phonemic Manipulation Task (Onset/Rime)

The Phonemic Manipulation Task was to take the onset of a first word (e.g., **R**ing) and the rime of a second word (e.g., h**AT**) to make a new real word, in this case **RAT**. Participants were instructed ahead of time about the difference between the onset (first sound) and the rime of a word, and were given many examples as well as several practice trials. All words used in this test were monosyllabic and, again, only pictures were used as stimuli (see **Figure 2**). Trials differed as to whether they could be completed correctly based on orthography alone, like the example above (called "shallow" trials), or could not (e.g., onset of '**B**ird' plus the rime of 't**OE**' makes a new word '**BO*W***'; called "deep"

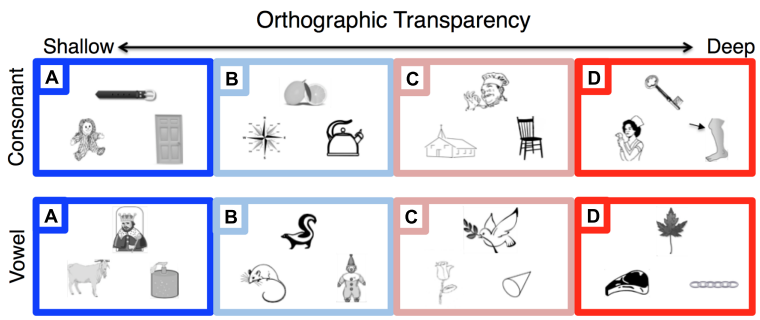


FIGURE 1 | Example stimuli for Phoneme Judgment Task. Participants had to pick the 'odd man out,' or which of the three pictures corresponded to an English name with a different first consonant sound (top row) or vowel sound (bottom row). For example, belt was the correct answer in the belt/doll/door triplet (top left). The orthographic transparency was manipulated in a graded manner such that orthographic information could help to accurately complete the Shallow (blue) conditions **(A,B)**, but would be uninformative or counter-productive in the Deep (red) conditions **(C,D)**. Shallow **(A)** trials were the most transparent, such that orthography alone could lead to the correct answer (e.g., first consonant: belt/doll/door; vowel: king/goat/soap). Shallow **(B)** trials could also be solved using more advanced orthographic knowledge (e.g., first consonant: lemon/compass/kettle; vowel: skunk, mouse, clown). Deep **(C)** trials did not give any orthographic cues, as all stimuli shared the same orthography of interest (e.g., first consonant: chef/church/chair; vowel: dove/rose/cone). Deep **(D)** trials gave counterproductive information such that using orthographic cues would systematically produce the wrong answer (e.g., first consonant: key/nurse/knee; vowel: leaf/steak/chain). The location of the odd man out was counterbalanced within a participant, but was placed at the top in each example above for clarity.

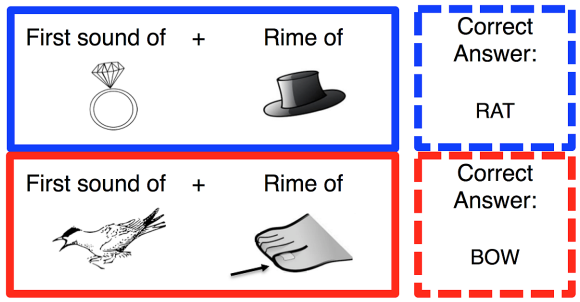


FIGURE 2 | Example stimuli in the Phonemic Manipulation Task. Participants were told to take the first sound from the left image and the rime of the right image to make a new word. The shallow trials (blue) were designed so that the task could be completed based on orthography alone (e.g., the first sound of 'Ring' + the rime of 'hAT' = RAT). The deep (red) trails were designed such that orthography alone could not be used to accurately complete the task (e.g., the first sound of 'Bird' + the rime of 'tOE' = BOW). The correct answer was not provided to participants as feedback, but is provided in the figure for clarity.

trials). Both shallow and deep trials were administered in the practice session. All subjects responded by typing their answer into the computer.

Results: Experiment 1

Phoneme Judgment Task

A 4 × 2 × 2 ANOVA was conducted with *orthographic transparency* (A, B, C, D) and *phoneme type* (consonant, vowel) as repeated measures, and *group* (deaf native signers, oral deaf) as a between subjects factor (see **Figure 3**). The main effect of orthographic transparency, $F(3,135) = 67.40$, $\eta^2 = 0.60$, $p < 0.001$, was significant in the predicted direction: the conditions that could be solved by transparent spelling alone were more accurate than those that required knowledge of the orthographic-to-phonological regularities, with the condition where an orthographic strategy would lead to consistently incorrect responses being the worst. There was a main effect of phoneme type, $F(1,45) = 22.13$, $\eta^2 = 0.33$, $p < 0.001$, such that responses in the vowel condition were more accurate than those in the consonant condition. Lastly, there was a main effect of group, $F(1,45) = 23.43$, $\eta^2 = 0.34$, $p < 0.001$, such that the oral deaf were more accurate than the deaf native signers. All three two-way interactions were significant. The orthographic transparency × group interaction was significant, $F(3,135) = 8.83$, $\eta^2 = 0.16$, $p < 0.001$, such that deaf native signers performance decreased more sharply as orthographic transparency diminishes than that of the oral deaf. The phoneme type × group interaction was significant, $F(1,45) = 6.00$, $\eta^2 = 0.12$, $p = 0.02$, such that the deaf native signers performed relatively worse on the first consonant condition, compared to the vowel condition, than did the oral deaf. Lastly, there was a significant orthographic transparency × phoneme type interaction, $F(3,135) = 9.24$, $\eta^2 = 0.17$, $p < 0.001$, such that the effect of orthographic transparency was more pronounced in the first consonant condition compared to the vowel condition. There was no significant three-way orthographic transparency × phoneme type × group interaction, $F(3,135) = 2.01$, $\eta^2 = 0.04$, $p = 0.12$.

Phonemic Manipulation Task

Data from the Phonemic Manipulation Task was entered into a 2 × 2 ANOVA with orthographic transparency (shallow, deep) as a repeated measure and group (oral deaf, deaf native signers) as a between subjects factor (see **Figure 4**). There was a significant main effect of orthographic transparency, $F(1,45) = 96.25$, $\eta^2 = 0.68$, $p < 0.001$, such that participants were less accurate in the deep condition where a transparent orthographic strategy could not be used successfully compared to the shallow condition. There was also a significant main effect of group, $F(1,45) = 41.86$, $\eta^2 = 0.48$, $p < 0.001$, such that the oral deaf had greater accuracy than deaf native signers. Lastly, there was a significant interaction between orthographic transparency and group, $F(1,45) = 38.63$,

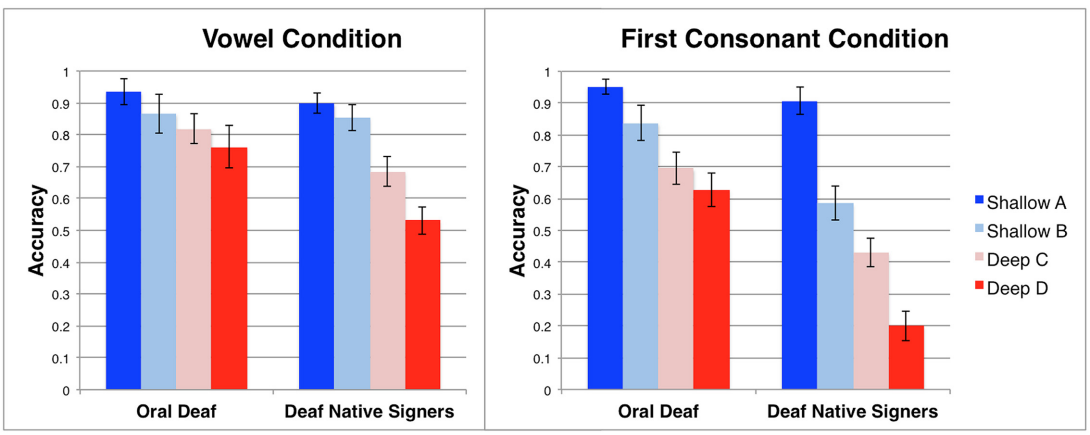


FIGURE 3 | Performance on Phoneme Judgment Task (Vowels and Consonants) across groups. The orthographic transparency was manipulated in a graded manner (see **Figure 1**) such that orthographic information could help to accurately complete the Shallow (blue) conditions (A,B), but would be uninformative or counter-productive in the Deep (red) conditions (C,D). Error bars represent SE of the mean.

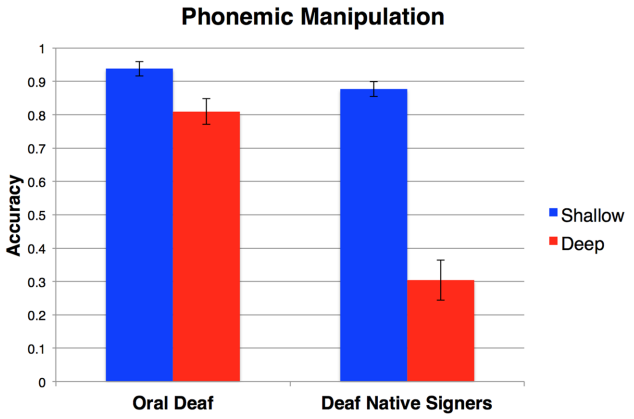


FIGURE 4 | Accuracy on the Phonemic Manipulation Task across groups. The shallow condition (blue) could be solved using an orthographic strategy, while the deep (red) condition required phonological knowledge above and beyond orthography. Error bars represent SE of the mean.

$\eta^2 = 0.46$, $p < 0.001$, such that deaf native signers performance decreased more sharply from shallow to deep than did the oral deaf performance.

For the separate group of hearing participants run to validate the tasks in Experiment 1, we confirm a significant effect of orthographic transparency in the Phoneme Judgment Task, the Phonemic Manipulation Task and when comparing the Phoneme Composite Scores (see Supplemental Information).

Experiment 1 Summary

Experiment 1 used two different tasks that systematically manipulated the extent to which orthographic information could be relied upon to access phonemic information. As expected, there was a strong effect of orthographic transparency on accuracy such that responses in shallow conditions were more accurate than in deep conditions. Although both deaf groups were sensitive to orthographic transparency, its impact was more pronounced in deaf native signers. This was the case for both the Phoneme Judgment Task and the Phonemic Manipulation Task. In terms of phoneme types, the vowel condition was easier overall than the consonant condition. Indeed, in the consonant condition of the Phoneme Judgment Task, performance in both deaf groups decreased sharply as orthography became less informative or counter-productive, and this effect was less pronounced in the vowel condition. One may speculate that this may reflect the fact that vowels tend to be more overtly enunciated on the lips (e.g., /e/ and /o/ are clearly differentiated on the lips), whereas many consonant distinctions are impossible to see on the lips (e.g., /ch/ vs. /sh/ or /g/ vs. /k/). Accordingly, greater accessibility through speechreading has been suggested to influence phonological knowledge in deaf populations in previous works (Erber, 1974; Walden et al., 2001).

Overall the main emerging pattern is that both deaf populations have a robust knowledge of orthographic regularities in English; however, deaf native signers show a greater reliance on visual orthographic information than the oral deaf when asked to complete English phonological tasks, at least when tested at the level of individual phonemes.

Experiment 2

The goal of Experiment 2 was to determine the best predictors of reading comprehension within each group, and compare how they may differ across the two deaf populations. Along with phonological knowledge, the contributions of memory skills that tap either phonological or semantic processing were also assessed in each group. Experiment 2 aims to determine how *useful* these skills may be in the service of reading comprehension in each of these deaf populations and whether group differences may emerge in best predictors. More specifically, we predict that oral deaf, with greater experience with spoken English, will make greater use of speech-based skills than deaf native signers (Lichtenstein, 1998).

A test of English reading comprehension was selected to evaluate reading skill, as many deaf adults, especially native signers, report that it is unnatural for them to read aloud. All participants completed the Peabody Individual Achievement Test-Revised: Reading Comprehension (Markwardt, 1989). This particular test is well tailored to deaf populations as it evaluates reading comprehension at the sentence level via non-verbal responses and has no speech production requirement (Morere, 2012). Participants were the same as in Experiment 1, meaning that the groups' performance on the TONI-3, a test of non-verbal spatial intelligence (Brown et al., 1997), did not significantly differ.

In addition to reading comprehension, measures known to be linked to reading comprehension skill were collected in order to assess if they differentially predicted reading comprehension across groups. These measures assessed knowledge of English phonology at different levels (Shallow and Deep Phoneme Composite Scores, Syllable Number Judgment, and Speechreading) and also different aspects of memory (serial recall span, primacy in a free recall span task).

Methods

Design and Procedure

Reading comprehension

The Peabody Individual Achievement Test-Revised: Reading Comprehension requires participants to read sentences one at a time and decide which of four pictures best matched the sentence just read. As the test progressed, the sentences increased in length, contained a greater number of clauses, and used less frequent vocabulary. Non-matching pictures were foils designed to represent erroneous interpretations that are based on expectations, and not on careful reading of the text. Thus, a reader must completely understand the grammar and vocabulary of the sentence in order to select the correct picture match. Instead of focusing on print-to-sound reading, as many reading tests do, this test focuses on lexical *and* syntactic knowledge of English. This test has been shown to be well suited to deaf populations (for a critique in hearing populations, see Keenan et al., 2006).

Phonological measures

Shallow and Deep Phoneme Composite Scores were derived from Experiment 1. In addition, performance on two other phonological tasks was collected. These tasks tapped larger units of English phonology, respectivively syllabic structure and sentence-level speechreading ability.

Phoneme Composite Scores

Accuracy on the Phoneme Judgment Task and the Phonemic Manipulation Task from Experiment 1 was collapsed across conditions to produce two composite scores. The first reflects performance in transparent conditions and was termed the ***Shallow Phoneme Composite Score***. It was derived from mean performance on the first two levels in the Phoneme Judgment Task (A, B) and from the shallow condition in the Phonemic Manipulation Task. The second reflects performance when spelling-to-sound correspondence is challenging, either because of the use of subtle featural differences (e.g., chef versus chair) or irregular orthography ('phone' shares a first sound with 'fence' and not 'paper'). It was named the ***Deep Phoneme Composite Score*** and is the mean performance in the Phoneme Judgment Task (C, D) and the deep condition in the Phonemic Manipulation Task.

Syllable Number Judgment Task

The Syllable Number Judgment Task also used a picture-based 'odd-man-out' paradigm. Participants were asked to select the item whose corresponding English name has a different number of syllables to the other two items. In order to prevent the use of word length as a strategy, words in each triad all contained the same number of letters and were either 5 or 6 letters long. All stimuli were picture-based. The odd man could either have more or fewer syllables than the other two items (e.g., **lemon**/clock/sheep or **glass**/table/paper).

Speechreading task

The speechreading task developed by Mohammed et al. (2003, 2006) was adapted to American English by using a native American English speaker to voice the sentences. Participants saw 15 spoken sentences (with no sound). After each sentence, participants had to select one from six pictures that best corresponded to the sentence just viewed. Picture foils were designed such that the observer must comprehend the whole sentence in order to answer correctly. For example, all six pictures that accompanied the sentence 'They were under the table' contained tables, three had more than one person, and one had a single person under a table, etc. Three practice sentences were given as preparation.

Short-term memory task – serial recall letter span

Separate lists of video stimuli of letters in English and in ASL were presented at a rate of 1 letter/sec. Visual ASL stimuli and audiovisual English stimuli were presented on the computer screen one at a time. ASL stimuli consisted of a native signer fingerspelling a list of letters and English stimuli consisted of a native speaker enunciating a list of letters in English. Lists ranged from 2 to 9 items in length, with two different lists at each length. The letters in the lists were the same as those used in Bavelier et al. (2008). Letters in both English and ASL were selected to be maximally dissimilar within each language in order to avoid phonological similarity effects (i.e., possible English written letters were: M, Y, S, L, R, K, H, G, P; ASL fingerspelled letters were: B, C, D, F, G, K, L, N, S). Participants were asked to repeat back each list in the precise order in which it was presented. The span was defined as the longest list length (L) recalled without mistakes before both list presentations in the next list length (L + 1) contained an error (e.g., if a participant recalled one list at length five correctly, but missed both lists at length six, their span would be five). Serial recall span was measured in each participants' preferred language (ASL for deaf native signers and English for oral deaf participants).

Free recall span

Participants were presented with lists of 16 words in English or in ASL, at the rate of 1 word every 5 s. Stimuli were videos of a native speaker or signer producing the list of 16 words, with a blank screen between each word. After viewing each list, they were required to immediately recall in their preferred language as many words as possible in any order. Each subject saw one list in each language and was told to try their best if it was not in their native language (e.g., spoken English for native signers or ASL for oral deaf). The items in each list were randomly assigned on a subject-by-subject basis from a list of 32 words, in order to avoid unplanned differences in word combinations that would lead one list to being 'easier' than the other. The lists used were roughly matched across groups, as much as possible with unequal sample sizes. Here we will only consider performance on the list in each participants' preferred language (ASL for deaf native signers and English for oral deaf). Measures of span, primacy and recency were derived from this data. Span was defined as the number of items recalled correctly (Rundus and Atkinson, 1970), primacy and recency scores were defined as the number of words recalled from among the first four (primacy) or last four (recency) items of the lists (Murdoch, 1962).

Results Experiment 2

Performance on Individual Tasks

Reading comprehension (PIAT grade-equivalent)

There was no main effect of group on reading comprehension scores, $t(45) = 0.44$, $d = 0.13$, $p = 0.66$.

Phonological Composite Scores

A 2 × 2 ANOVA on the composite accuracy scores with composite score type (shallow, deep) as a repeated measures and group (deaf native signer, oral deaf) as a between subjects factor revealed, as expected given the previous analyses, main effects of composite score type, $F(1,45) = 181.83$, $\eta^2 = 0.80$, $p < 0.001$, and group, $F(1,45) = 33.00$, $\eta^2 = 0.42$, $p < 0.001$. There was also a significant interaction, $F(1,45) = 31.43$, $\eta^2 = 0.41$, $p < 0.001$, such that the effect of orthographic transparency (deep vs. shallow) was greater for deaf native signers, $t(25) = 13.70$, $d = 5.48$, $p < 0.001$, than it was for the oral deaf, $t(20) = 5.61$, $d = 2.51$, $p < 0.001$ (**Figure 5**).

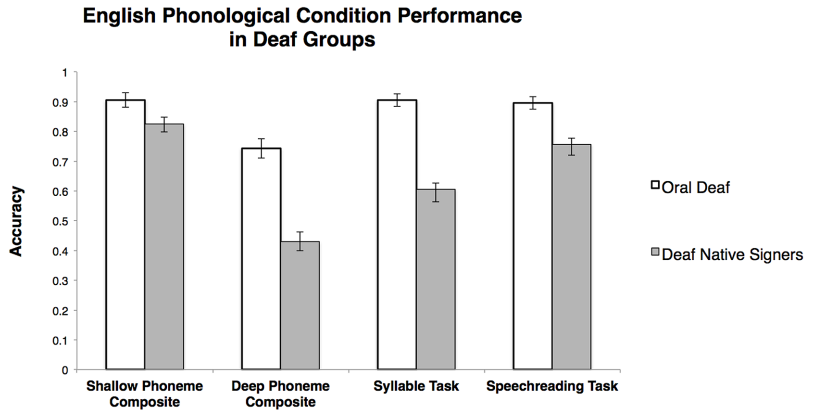


FIGURE 5 | English Phonological Task Performance. Error bars represent the SE of the mean.

Syllable Number Judgment Task

There was a significant effect of group on accuracy in the Syllable Number Judgment Task, $t(45) = 5.93$, $d = 1.77$, $p < 0.001$, such that the oral deaf group performed significantly better than the deaf native signer group (**Figure 5**).

Speechreading Task

There was a significant effect of group on the speechreading task, $t(45) = 3.09$, $d = 0.92$, $p < 0.001$, such that the oral deaf group performed significantly better than the deaf native signer group (**Figure 5**).

Serial Recall Memory

The serial recall spans in deaf native signers and the oral deaf were comparable, $t(45) = 0.92$, $d = 0.27$, $p = 0.37$, and in the range of 5 ± 1 (**Figure 6**), as expected from the existing literature (Boutla et al., 2002, 2004; Koo et al., 2008).

Free recall memory

Free recall memory was measured in ASL and in English for each participant. However, here we only include performance in each participant's preferred language (English for the oral deaf; ASL for deaf native signers). Free recall span was defined as the total number of accurately recalled words from the list. There was no main effect of group, $t(45) = 1.67$, $d = 0.50$, $p = 0.10$. Analyses of the primacy and recency effects also revealed no main effects of group: primacy, $t(45) = 1.07$, $d = 0.32$, $p = 0.29$, and recency, $t(45) = 0.59$, $d = 0.18$, $p = 0.55$ (**Figure 6**)[3].

A key distinction for our study is that serial recall and primacy free recall tap into different memory processes. Accordingly, these two measures show little correlation in the deaf participants [$r(45) = 0.143$; $p = 0.34$].

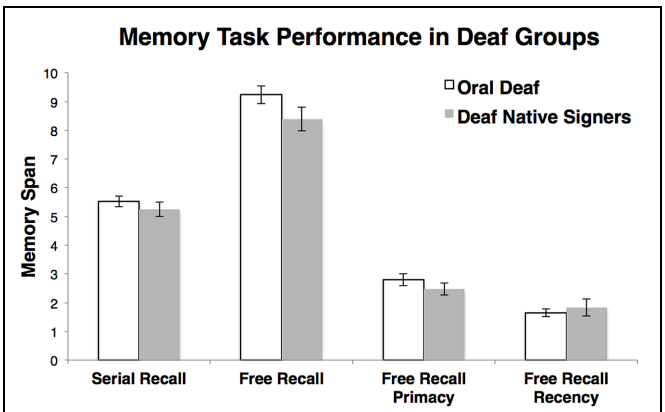


FIGURE 6 | Memory task performance. Error bars represent the SE of the mean.

Predictors of Reading Comprehension

The main question of interest concerns the variables that best predict reading comprehension and whether they differ between the two deaf populations. We first present an analysis of how reading predictors may differ across groups and then consider the impact of the different predictors within each group.

Group comparisons

Regression analyses were computed using R (R Development Core Team, 2010) with grade-equivalent PIAT scores as the dependent variable. We first removed all variance in PIAT scores attributable to non-verbal IQ as well as unaided dB loss in both groups, by regressing PIAT scores against TONI-3 scores and the unaided dB loss in the better ear. All further analyses were then performed on the residuals of this regression. Missing data was replaced with the mean, but whether or not missing non-verbal IQ or dB loss data was excluded pairwise or replaced with the mean, the significance levels of the models reported below did not change. Neither non-verbal IQ nor dB loss accounted for a significant amount of variance in any of the models.

First, in order to assess whether the predictors of reading comprehension were significantly different across the two deaf groups, two types of regression models were created. Model 1

[3]The stimulus list order was not available for three native signer participants due to a technical malfunction, and their primacy and recency scores were not possible to calculate. Their data was replaced with the group mean for native signers (Primacy = 2.47, Recency = 1.83).

was a main effect model, with eight predictor variables: Shallow Phoneme Composite Score, Deep Phoneme Composite Score, Syllable Number Judgment Task, Speechreading, Serial Recall, Free Recall Primacy, Free Recall Recency, and group (oral deaf, deaf native signer). Models 2_{a-g} separately added the interaction terms between group and the remaining seven predictors in a stepwise manner. A significant group × predictor interaction term would demonstrate a different level of importance of that given predictor for one group compared to the other. On its own, Model 1 was a significant predictor of reading performance [adjusted $R^2 = 0.33$; $F(8,36) = 3.67$, $p = 0.003$] indicating that together the eight predictors (including group) accounted for a significant amount of variance in reading comprehension across all deaf participants. Interestingly, the group × free recall primacy interaction was the only significant interaction term: $F(1,35) = 11.59$, $p = 0.002$ [Model 2: adjusted $R^2 = 0.48$; $F(9,35) = 5.51$, $p < 0.001$]. This demonstrates that the free recall primacy measure differentially affects reading comprehension in deaf native signers and oral deaf participants. As can be seen in **Figure 7**, free recall primacy was a better predictor of reading comprehension for deaf native signers than it was for the oral deaf.

There was a significant positive correlation between Free Recall Primacy and Reading Comprehension in the deaf native signers, $R^2 = 0.21$, $p = 0.02$, whereas there was no correlation in the oral deaf, $R^2 = 0.01$, $p = 0.67$. This analysis supports the hypothesis that determinants of reading comprehension are different for oral deaf and deaf native signers. To better understand the main determinants of reading comprehension in each population, each group was considered separately.

Individual group partial correlations

To confirm and elaborate on the results of the combined regressions above, partial correlations were separately computed for each group between reading comprehension (having removed variance due to TONI and hearing loss) and the remaining seven predictors: Shallow Phonological Composite Score, Deep Phonological Composite Score, Syllable Number Judgment Task, Speechreading, Serial Recall, Free Recall Primacy, and Free Recall Recency.

The strongest correlations with reading comprehension for the oral deaf were measures of English phonological knowledge, independent of orthographic knowledge. The Deep Phonological score, $r(18) = 0.66$, $p = 0.003$, as well as serial recall span, $r(18) = 0.50$, $p = 0.04$ correlated highly with reading comprehension. None of the other factors were significantly correlated with reading comprehension (all $ps > 0.12$). In stark contrast to the oral deaf, for the deaf native signers the Free Recall Primacy measure, $r(22) = 0.41$, $p = 0.04$, and the Shallow Phonological score, $r(18) = 0.52$, $p = 0.009$, were the only measures that significantly correlated with reading comprehension.

Discussion

This study compared determinants of reading in two distinct deaf populations with marked differences in language experience. The two deaf groups were selected to differ in their language experience, by recruiting either deaf native signers or oral deaf individuals. Both groups were exposed to a natural language in early childhood, but that language and ongoing language experience was signed in the case of deaf native signers and spoken in the case of the oral deaf. Importantly, these two groups had similar reading comprehension scores, as well as similar performance on general cognitive measures such as non-verbal IQ and free and serial memory recall. However, these two groups differed in what best predicted their reading comprehension scores. Whereas the reading comprehension of the oral deaf was best predicted by both deep phonological knowledge and serial recall span, deaf native signers' reading comprehension was best predicted by their performance on the free recall task. In particular, reading comprehension in deaf native signers showed a significant correlation with the primacy component

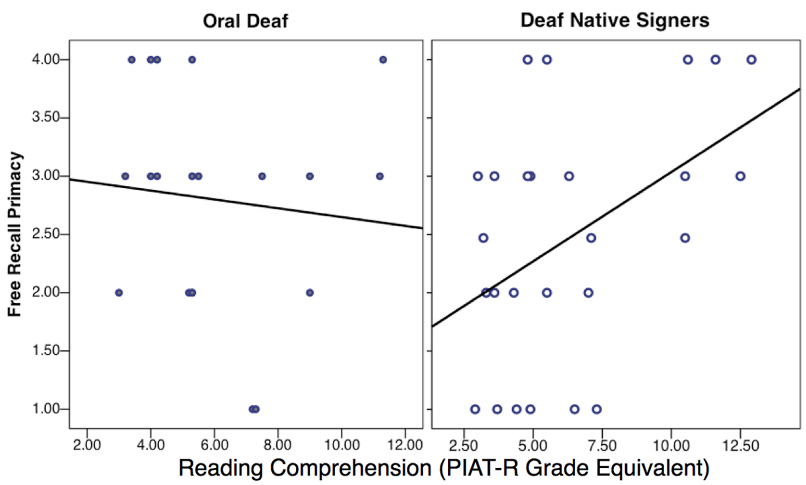


FIGURE 7 | Regression Plots of the Effect of Free Recall Primacy on Reading Comprehension in Deaf Groups.

of the free recall span, associated with short-term memory consolidation (Bayliss et al., 2015) and semantic coding (Craik and Lockhart, 1972; Martin and Saffran, 1997). More specifically, the link between reading comprehension and the primacy effect in the deaf native signers mirrors that reported by Bauer and Emhert (1984) who found that differences in the primacy effect, compared to the recency effect, better discriminated non-disabled from disabled readers.

English Phonological Knowledge in Deaf Individuals

There still remains outstanding questions about whether deaf readers, especially oral ones, have qualitatively similar English phonological knowledge to that of hearing individuals. There are different ways that one can acquire English phonological knowledge. It can be acquired from auditory information (such as hearing the difference between a voiced and voiceless glottal stop (/g/ and /k/), from articulatory information as when speaking and speechreading by observing the movement of the lips and mouth, or from a tutored visual experience such as is the case with Cued Speech (LaSasso et al., 2003), or even from orthography during reading in alphabetic language like English. The extant literature on Cued Speech for example makes it clear that such communication training enhances awareness of phonological knowledge for the trained spoken language (Alegria and Lechat, 2005). The resulting phonological knowledge has been shown to be comparable to that of both oral deaf and hearing individuals (Koo et al., 2008) and to facilitate reading skills (Colin et al., 2007). In the present study, our two deaf populations share the fact that they were born profoundly deaf, which makes them different from hearing individuals, but they also differ amongst themselves in their language experience, residual hearing, and use of hearing aids or CIs. Indeed, oral deaf individuals are more likely to attain information from articulation, visual speechreading experience, or aided residual hearing, whereas native signers are most likely acquiring phonological information solely through visual experiences such as reading and limited speechreading. These differences are reflected in the performance of these two groups on the phonological tasks presented in this work. For example, native signers were more likely to perform poorly than the oral deaf in the deep phonological conditions, where orthography was uninformative or misleading.

The current study also provides some insights for cross-linguistic studies of phonological skill in deafness. In addition to the importance of carefully considering population characteristics, we demonstrate that the nature of the orthographic-phonological mapping of a written language may also be important. In light of these considerations, the lack of an effect of language experience (speech versus sign) on phonological awareness in a study conducted in Hebrew is worth considering (Miller, 1997). Hebrew has a relatively simple mapping between orthography and sound and has multiple letters that map onto the same phonemes, like English. Interestingly, conditions that required that type of knowledge (e.g., knowing that when deciding the odd man out between 'c', 'k' and 'p', that 'c' and 'k' sometimes sound the same) did not reveal major differences between oral and signing deaf participants in the current work. Yet, clearly oral deaf subjects differ from deaf native signers in their knowledge of English phonology. Such differences may not be as easily detectable in a transparent language such as Hebrew.

Phonological Awareness and Reading Comprehension in Deaf Individuals

The current study also aimed to address concerns about the link between phonological awareness measures and reading scores in two different deaf populations. For the oral deaf, it was the variance in tasks that require English phonological knowledge, above and beyond orthographic knowledge, that best predicted reading. In contrast, for the deaf native signers, in addition to free recall being a good predictor, the measure of phonological skills that best predicted reading was one that could be solved by visual information alone or by conceptual knowledge about spelling. The inclusion of deaf groups with different language experience makes it clear that not all deaf populations possess the same phonological knowledge of English. The use of tasks that systematically manipulated the relationships between phonology and orthography was crucial in being able to draw this conclusion. Our study may explain some of the conflicting reports in the literature (Mayberry et al., 2011) since past studies have included populations that varied significantly in their language experience, all encompassed under the term "deaf." Furthermore, our study confirms the need to avoid phonological tasks that confound orthographic and phonological knowledge (McQuarrie and Parrila, 2009). The results highlight the importance of a detailed analysis of both the characteristics of the language/script to be read and the population of deaf individuals studied.

The shallow phonological score essentially measures orthographic knowledge or familiarity with spelling, and the usefulness of such information in inferring the phonological structure of English. We did find that it accounted for a significant amount of the variance in reading comprehension in deaf native signers. This score could be linked with single word processing and identification, but without access to more detailed statistics on the participants' reading habits it is also possible that the shallow phonological score reflects exposure to print, being in a sense an indirect measure of reading skill. Indeed, greater exposure to print could lead to greater orthographic knowledge and better word identification skills, which could in turn lead to overall greater reading skill and comprehension. Further experiments are necessary to clarify the relationship between performance in our shallow phonological conditions, the use of orthography in phonological tasks, and reading comprehension in the deaf.

Reading Comprehension and Free Recall Memory in Deaf Native Signers

Finally and probably most importantly, the present work indicates that memory processes associated with the free recall task may provide an alternative route for supporting reading

in deaf native signers. Primacy scores in the free recall task, associated with semantic processing, was the one predictor that differentially predicted reading comprehension in deaf native signers and the oral deaf. Studies that recruit deaf participants without considering their language experience are likely to encompass only a very small percentage of deaf native signers given their low prevalence, resulting in an over-emphasis on the role of English phonological skills compared to semantic-based memory skill in deaf reading. This may explain why our study is the first one to highlight this link, despite a strong relationship between free recall and reading comprehension in our deaf native signing participants[4].

These results need to be situated in the larger picture of what we know about reading processes. A first intriguing issue concerns what it may mean for a free recall task tested in American Sign Language to be a good predictor of comprehension of English text in deaf native signers. Due to the connection in the literature between free recall, with a focus on the primacy effect, and semantic processing, one interpretation could be that deaf native signers rely to a greater extent on processing of semantic information at both the word level and the sentence level in the service of reading comprehension. For example, semantic processing is necessary to maintain coherence, hold information online in memory, and make appropriate connections within and between phrase structures in order to comprehend a text. Deficits in semantic processing have been linked to poor comprehension skill (Nation and Snowling, 1998b; Hagtvet, 2003; Cain and Oakhill, 2006). It is possible that enhanced semantic processing, or at least a greater *reliance* on semantic processing (Sinatra et al., 1984; Nation and Snowling, 1998a), may help compensate for deficient phonological skills. Accordingly, top–down semantic influences on deaf readers, such as prior knowledge or context (Kelly, 1995; Jackson et al., 1997) have been shown to be significant predictors of passage comprehension, which is consistent with our current findings. Since ASL grammar is quite different from that of English, deaf native signers not only have to identify words in another language, but they need to understand the syntactic rules that connect them. Yurkowski and Ewoldt (1986) proposed that semantic information maybe crucial in helping with complex syntactic processing.

Another interesting perspective is that deaf native signers are actually bilingual (bi-modal) readers and thus reading their second language when faced with English text (Chamberlain and Mayberry, 2008). Our findings are consistent with the ideas put forth by Ullman (2001, 2005) which suggest that second language learners rely more on lexical memory, supported by the declarative memory system. For example, several studies indicate that non-proficient hearing speakers while reading in their second language differ from first language readers on measures of integration, recognition of aspects of text structure, use of general knowledge, and personal experience, as well as in paying attention to 'broader phrases' and keeping the meaning of the passages in mind during reading (Carrell, 1989; Fitzgerald, 1995; Jun Zhang, 2001). Primacy in free recall, also thought to be a measure linked to semantic processing (Craik and Tulving, 1975; Bellezza et al., 1976; Waters and Waters, 1976), could be related to such cognitive skills that highlight the role of recognition and integration of memory representations over broader linguistic units.

Conclusion

In sum, the present work clarifies the nature of English phonological knowledge in two distinct deaf populations: deaf native signers and the oral deaf. It highlights the importance of considering language experience when evaluating determinants of reading in deaf participants. It also reveals for the first time a potential complementary route to literacy – semantic-based memory – that does not depend upon English phonological skills. It will be for future research to assess precisely how greater reliance on semantic processing may foster good text comprehension, even in the face of poor phonological skills.

Acknowledgments

We would like to thank all of the subjects recruited from the National Technical Institute of the Deaf at the Rochester Institute of Technology, Rochester, NY, USA. We would also like to thank P. Clark, B. McDonald, and A. Hauser for their invaluable interpreting services. This research was supported by the National Institutes of Health (DC04418 to DB) and the Charles A. Dana Foundation (to DB and MD).

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01153

[4]In order to ensure that our findings were not a result of a general trend in all deaf readers, regardless of language background, we combined the data from both deaf groups and created high and low median split groups based on their PIAT reading comprehension scores. Both PIAT skill groups were similarly represented by oral deaf and native signers (LowPIAT group contained 11 oral deaf and 12 native signers; HighPIAT group contained 10 oral deaf and 14 native signers). Using these newly defined groups, there were no significant differences in the deep phonological measure ($M_{\text{HighPIAT}} = 0.61$, SDHighPIAT = 0.23; MLowPIAT = 0.53, SDLowPIAT = 0.19, $p = 0.23$) or shallow phonological measure (MHighPIAT = 0.89, SDHighPIAT = 0.10; MLowPIAT = 0.83, SDLowPIAT = 0.15, $p = 0.11$).
While this may seem surprising given the broad patterns in the literature at large, this result highlights the importance and consequences of combining data from deaf individuals with distinct language backgrounds. A distinctive feature of our study is to have carefully selected groups that have largely homogenous within-group language background and distinct between-group language background. By considering good/bad readers irrespective of language background, we are essentially diluting each of our deaf group's effects. Accordingly, there were also no significant differences in the free recall primacy scores when considering good/poor readers over the whole deaf group ($M_{\text{HighPIAT}} = 2.69$, SDHighPIAT = 1.05; MLowPIAT = 2.54, SDLowPIAT = 0.99, $p = 0.62$). Such results highlight the importance of considering separately the "oral" deaf and the "deaf signers" group.

References

Alegria, J., and Lechat, J. (2005). Phonological processing in deaf children: when lipreading and cues are incongruent. *J. Deaf Stud. Deaf Educ.* 10, 122–133. doi: 10.1093/deafed/eni013

Aparicio, M., Gounot, D., Demont, E., and Metz-Lutz, M. N. (2007). Phonological processing in relation to reading: an fMRI study in deaf readers. *Neuroimage* 35, 1303–1316. doi: 10.1016/j.neuroimage.2006.12.046

Baddeley, A. D., Lewis, V., and Vallar, G. (1984). Exploring the articulatory loop. *Q. J. Exp. Psychol.* 36A, 233–252. doi: 10.1080/14640748408402157

Bauer, R. H., and Emhert, J. (1984). Information processing in reading-disabled and nondisabled children. *J. Exp. Child Psychol.* 37, 271–281. doi: 10.1016/0022-0965(84)90005-5

Bavelier, D., Newport, E. L., Hall, M. L., lla, T., and Boutla, M. (2008). Ordered short-term memory differs in signers and speakers: implications for models of short-term memory. *Cognition* 107, 433–459. doi: 10.1016/j.cognition.2007.10.012

Bayliss, D. M., Bogdanovs, J., and Jarrold, C. (2015). Consolidating working memory: distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *J. Mem. Lang.* 81, 34–50. doi: 10.1016/j.jml.2014.12.004

Bélanger, N. N., Baum, S. R., and Mayberry, R. I. (2012a). Reading difficulties in adult deaf readers of French: phonological codes, not guilty! *Sci. Stud. Read.* 16, 263–285. doi: 10.1080/10888438.2011.568555

Bélanger, N. N., Slattery, T. J., Mayberry, R. I., and Rayner, K. (2012b). Skilled deaf readers have an enhanced perceptual span in reading. *Psychol. Sci.* 23, 816–823. doi: 10.1177/0956797611435130

Bélanger, N. N., Mayberry, R. I., and Rayner, K. (2013). Orthographic and phonological preview benefits: parafoveal processing in skilled and less-skilled deaf readers. *Q. J. Exp. Psychol.* 66, 2237–2252. doi: 10.1080/17470218.2013.780085

Bellezza, F. S., Richards, D. L., and Geiselman, R. E. (1976). Semantic processing and organization in free recall. *Mem. Cogn.* 4, 415–421. doi: 10.3758/BF03213198

Bellugi, U., Klima, E., and Siple, P. (1975). Remembering in Signs. *Cognition* 3, 93–125. doi: 10.1016/0010-0277(74)90015-8

Bhatarah, P., Ward, G., Smith, J., and Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Mem. Cogn.* 37, 689–713. doi: 10.3758/MC.37.5.689

Boutla, M., Supalla, T., and Bavelier, D. (2002). What can American Sign Language tell us about capacity limit in working memory? *Paper Presented at the Cognitive Neuroscience Society*, San Francisco, CA.

Boutla, M., Supalla, T., Newport, L., and Bavelier, D. (2004). Short-term memory span: insights from sign language. *Nat. Neurosci.* 7, 997–1002. doi: 10.1038/nn1298

Brown, L. (2003). *Test of Nonverbal Intelligence Handbook of Nonverbal Assessment.* (Berlin: Springer), 191–221. doi: 10.1007/978-1-4615-0153-4_10

Brown, L., Sherbenou, R. J., and Johnsen, S. K. (1997). *Test of Nonverbal Intelligence*, 3rd Edn. Austin, TX: PRO-ED.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Dev. Psychol.* 28, 874. doi: 10.1037/0012-1649.28.5.874

Burgess, N., and Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychol. Rev.* 106, 551. doi: 10.1037/0033-295X.106.3.551

Cain, K., and Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *Br. J. Educ. Psychol.* 76, 683–696. doi: 10.1348/000709905X67610

Carrell, P. L. (1989). Metacognitive awareness and second language reading. *Mod. Lang. J.* 73, 121–134. doi: 10.1111/j.1540-4781.1989.tb02534.x

Chamberlain, C., and Mayberry, R. I. (2000). "Theorizing about the relation between American Sign Language and reading," in *Language Acquisition by Eye*, eds C. Chamberlain, J. P. Morford, and R. I. Mayberry (ahwah, NJ: Earlbaum), 221–259.

Chamberlain, C., and Mayberry, R. I. (2008). American Sign Language syntactic and narrative comprehension in skilled and less skilled readers: bilingual and bimodal evidence for the linguistic basis of reading. *Appl. Psycholinguist.* 29, 367–388. doi: 10.1017/S014271640808017X

Colin, S., Magnan, A., Ecalle, J., and Leybaert, J. (2007). Relation between deaf children's phonological skills in kindergarten and word recognition performance in first grade. *J. Child Psychol. Psychiatry* 48, 139–146. doi: 10.1111/j.1469-7610.2006.01700.x

Coltheart, V., Laxon, V., Rickard, M., and Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *J. Exp. Psychol.* 14, 387. doi: 10.1037/0278-7393.14.3.387

Conrad, R. (1972). Short-term memory in the Deaf: a test for speech coding. *Br. J. Psychol.* 63, 173–180. doi: 10.1111/j.2044-8295.1972.tb02097.x

Cornwall, A. (1992). The relationship of phonological awareness, rapid naming, and verbal memory to severe reading and spelling disability. *J. Learn. Disabil.* 25, 532–538. doi: 10.1177/002221949202500808

Craik, F. I. M., and Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *J. Verbal Learning Verbal Behav.* 11, 671–684. doi: 10.1016/S0022-5371(72)80001-X

Craik, F. I. M., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol.* 104, 268–294. doi: 10.1037/0096-3445.104.3.268

Dallago, M. L. L., and Moely, B. E. (1980). Free recall in boys of normal and poor reading levels as a function of task manipulations. *J. Exp. Child Psychol.* 30, 62–78. doi: 10.1016/0022-0965(80)90075-2

Dyer, A., MacSweeney, M., Szczerbinski, M., Green, L., and Campbell, R. (2003). Predictors of reading delay in deaf adolescents: the relative contributions of rapid automatized naming speed and phonological awareness and decoding. *J. Deaf Stud. Deaf Educ.* 8, 215. doi: 10.1093/deafed/eng012

Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffiro, T. A., et al. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron* 44, 411–422. doi: 10.1016/j.neuron.2004.10.019

Elbro, C., Nielsen, I., and Petersen, D. K. (1994). Dyslexia in adults: evidence for deficits in non-word reading and in the phonological representation of lexical items. *Ann. Dyslexia* 44, 203–226. doi: 10.1007/BF02648162

Erber, N. P. (1974). Visual perception of speech by deaf children: recent developments and continuing needs. *J. Speech Hear. Disord.* 39, 178–185. doi: 10.1044/jshd.3902.178

Figueras, B., Edwards, L., and Langdon, D. (2008). Executive function and language in deaf children. *J. Deaf Stud. Deaf Educ.* 13, 362–377. doi: 10.1093/deafed/enm067

Fitzgerald, J. (1995). English-as-a-second-language learners, Äô cognitive reading processes: a review of research in the United States. *Rev. Educ. Res.* 65, 145–190. doi: 10.3102/00346543065002145

Fletcher, C. R. (1986). Strategies for the allocation of short-term memory during comprehension. *J. Mem. Lang.* 25, 43–58. doi: 10.1016/0749-596X(86)90020-3

Gabrieli, J. D. E. (2009). Dyslexia: a new synergy between education and cognitive neuroscience. *Science* 325, 280. doi: 10.1126/science.1171999

Gallaudet Research Institute. (2005). Available at: https://research.gallaudet.edu/Demographics/

Gathercole, S. E., Willis, C., and Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: reading and vocabulary development in children. *Br. J. Psychol.* 82, 387–406. doi: 10.1111/j.2044-8295.1991.tb02407.x

Geraci, C., Gozzi, M., Papagno, C., and Cecchetto, C. (2008). How grammar can cope with limited short-term memory: simultaneity and seriality in sign languages. *Cognition* 106, 780–804. doi: 10.1016/j.cognition.2007.04.014

Goldin Meadow, S., and Mayberry, R. I. (2001). How do profoundly deaf children learn to read? *Learn. Disabil. Res. Pract.* 16, 222–229. doi: 10.1111/0938-8982.00022

Gozzi, M., Geraci, C., Cecchetto, C., Perugini, M., and Papagno, C. (2011). Looking for an explanation for the low sign span: is order involved? *J. Deaf Stud. Deaf Educ.* 16, 101–107. doi: 10.1093/deafed/enq035

Hagtvet, B. E. (2003). Listening comprehension and reading comprehension in poor decoders: evidence for the importance of syntactic and semantic skills as well as phonological skills. *Read. Writ.* 16, 505–539. doi: 10.1023/A:1025521722900

Hall, M. L., and Bavelier, D. (2010). "Working memory, deafness, and sign language," in *The Oxford Handbook of Deaf Studies, Language, and Education*,

Vol. 2, eds M. Marschark, P. E. Spencer, and P. E. Nathan (New York: Oxford University Press), 458–472.

Hammill, D. D. (1987). *Test of Adolescent Language (TOAL-2): A Multidimensional Approach to Assessment: Pro-Ed*, Vol. 2. (Austin, TX: Pro-Ed Inc.).

Hanson, V. L., and Fowler, C. A. (1987). Phonological coding in word reading: evidence from hearing and deaf readers. *Mem. Cogn.* 15, 199–207. doi: 10.3758/BF03197717

Hanson, V. L., and McGarr, N. S. (1989). Rhyme generation by deaf adults. *J. Speech Hear. Res.* 32, 2. doi: 10.1044/jshr.3201.02

Hatcher, P. J., Hulme, C., and Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: the phonological linkage hypothesis. *Child Dev.* 65, 41–57. doi: 10.2307/1131364

Hauser, P. C., Lukomski, J., and Hillman, T. (2008a). "Development of deaf and hard-of-hearing students' executive function," in *Deaf Cognition*, eds M. H. Marschark and P. C. Hauser (Oxford: Oxford University Press), 286–308.

Hauser, P. C., Paludnevičiene, R., Supalla, T., and Bavelier, D. (2008b). "American sign language-sentence reproduction test: development and implications," in *Sign Language: Spinning and Unraveling the Past, Present and Future,* ed. R. M. D. Quadros (Petropolis: Editora Arara Azul), 160–172.

Hirshorn, E. A., Fernandez, N. M., and Bavelier, D. (2012). Routes to short-term memory indexing: lessons from deaf native users of American Sign Language. *Cogn. Neuropsychol.* 29, 85–103. doi: 10.1080/02643294.2012.704354

Hogan, T. P., Catts, H. W., and Little, T. D. (2005). The relationship between phonological awareness and reading: implications for the assessment of phonological awareness. *Lang. Speech Hear. Serv. Sch.* 36, 285–293. doi: 10.1044/0161-1461(2005/029)

Høien, T., Lundberg, I., Stanovich, K. E., and Bjaalid, I.-K. (1995). Components of phonological awareness. *Read. Writ.* 7, 171–188. doi: 10.1007/BF01027184

Hyde, T. S., and Jenkins, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks1. *J. Verbal Learning Verbal Behav.* 12, 471–480. doi: 10.1016/S0022-5371(73)80027-1

Jackson, D., Paul, P., and Smith, J. (1997). Prior knowledge and reading comprehension ability of deaf adolescents. *J. Deaf Stud. Deaf Educ.* 2, 172. doi: 10.1093/oxfordjournals.deafed.a014323

Jolicœur, P., and Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cogn. Psychol.* 36, 138–202. doi: 10.1006/cogp.1998.0684

Jun Zhang, L. (2001). Awareness in reading: EFL students' metacognitive knowledge of reading strategies in an acquisition-poor environment. *Lang. Awareness* 10, 268–288. doi: 10.1080/09658410108667039

Keenan, J. M., Betjemann, R. S., Wadsworth, S. J., DeFries, J. C., and Olson, R. K. (2006). Genetic and environmental influences on reading and listening comprehension. *J. Res. Read.* 29, 75–91. doi: 10.1111/j.1467-9817.2006.00293.x

Kelly, L. P. (1995). Processing of bottom-up and top-down information by skilled and average deaf readers and implications for whole language instruction. *Except. Child.* 61, 318–334.

Koo, D., Crain, K., LaSasso, C., and Eden, G. F. (2008). Phonological awareness and short term memory in hearing and deaf individuals of different communication backgrounds. *Ann. N. Y. Acad Sci.* 1145, 83–99. doi: 10.1196/annals.1416.025

Landi, N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Read. Writ.* 3, 701–717. doi: 10.1007/s11145-009-9180-z

LaSasso, C., Crain, K., and Leybaert, J. (2003). Rhyme generation in deaf students: the effect of exposure to cued speech. *J. Deaf Stud. Deaf Educ.* 8, 250–270. doi: 10.1093/deafed/eng014

Lee, J. F. (1986). On the use of the recall task to measure L2 reading comprehension. *Stud. Second Lang. Acquis.* 8, 201–212. doi: 10.1017/S0272263100006082

Lichtenstein, E. H. (1998). The relationships between reading processes and English skills of deaf college students. *J. Deaf Stud. Deaf Educ.* 3, 80. doi: 10.1093/oxfordjournals.deafed.a014348

Luetke-Stahlman, B., and Nielsen, D. C. (2003). The contribution of phonological awareness and receptive and expressive English to the reading ability of deaf students with varying degrees of exposure to accurate English. *J. Deaf Stud. Deaf Educ.* 8, 464. doi: 10.1093/deafed/eng028

MacSweeney, M., Brammer, M. J., Waters, D., and Goswami, U. (2009). Enhanced activation of the left inferior frontal gyrus in deaf and dyslexic adults during rhyming. *Brain* 132, 1928–1940. doi: 10.1093/brain/awp129

MacSweeney, M., Waters, D., Brammer, M. J., Woll, B., and Goswami, U. (2008). Phonological processing in deaf signers and the impact of age of first language acquisition. *Neuroimage* 40, 1369–1379. doi: 10.1016/j.neuroimage.2007.12.047

Markwardt, F. C. (1989). *Peabody Individual Achievement Test-Revised: PIAT-R.* Circle Pines, MN: American Guidance Service.

Marshall, C. R., Mann, W., and Morgan, G. (2011). Short-term memory in signed languages: not just a disadvantage for serial recall. *Front. Psychol.* 2:102. doi: 10.3389/fpsyg.2011.00102

Martin, N., and Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: evidence from associations and dissociations. *Cogn. Neuropsychol.* 21, 213–228. doi: 10.1080/02643290342000447

Martin, N., and Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: evidence for common underlying processes. *Cogn. Neuropsychol.* 14, 641–682. doi: 10.1080/026432997381402

Mayberry, R. I. (1993). First-language acquisition after childhood differs from second-language acquisition: the case of American Sign Language. *J. Speech Lang. Hear. Res.* 36, 1258. doi: 10.1044/jshr.3606.1258

Mayberry, R. I., del Giudice, A. A., and Lieberman, A. M. (2011). Reading achievement in relation to phonological coding and awareness in deaf readers: a meta-analysis. *J. Deaf Stud. Deaf Educ.* 16, 164–188. doi: 10.1093/deafed/enq049

McCardle, P., Scarborough, H. S., and Catts, H. W. (2001). Predicting, explaining, and preventing children's reading difficulties. *Learn. Disabil. Res. Pract.* 16, 230–239. doi: 10.1111/0938-8982.00023

McDougall, S., Hulme, C., Ellis, A., and Monk, A. (1994). Learning to read: the role of short-term memory and phonological skills. *J. Exp. Child Psychol.* 58, 112–133. doi: 10.1006/jecp.1994.1028

McQuarrie, L., and Parrila, R. (2009). Phonological representations in deaf children: rethinking the functional equivalence hypothesis. *J. Deaf Stud. Deaf Educ.* 14, 137. doi: 10.1093/deafed/enn025

Melby-Lervåg, M., and Hulme, C. (2010). Serial and free recall in children can be improved by training evidence for the importance of phonological and semantic representations in immediate memory tasks. *Psychol. Sci.* 21, 1694–1700. doi: 10.1177/0956797610385355

Miller, P. (1997). The effect of communication mode on the development of phonemic awareness in prelingually deaf students. *J. Speech Lang. Hear. Res.* 40, 1151. doi: 10.1044/jslhr.4005.1151

Miller, P., and Clark, M. D. (2011). Phonemic awareness is not necessary to become a skilled deaf reader. *J. Dev. Phys. Disabil.* 23, 459–476. doi: 10.1007/s10882-011-9246-0

Mohammed, T., Campbell, R., Macsweeney, M., Barry, F., and Coleman, M. (2006). Speechreading and its association with reading among deaf, hearing and dyslexic individuals. *Clin. Linguist. Phon.* 20, 621–630. doi: 10.1080/02699200500266745

Mohammed, T. E., MacSweeney, M., and Campbell, R. (2003). "Developing the TAS: individual differences in silent speechreading, reading and phonological awareness in deaf and hearing speechreaders," in *Proceedings of the AVSP 2003 – International Conference on Audio-Visual Speech Processing*, Saint-Jorioz, 49–54.

Morere, D. A. (2012). "Measures of reading achievement," in *Assessing Literacy in Deaf Individuals,* eds T. Allen and D. A. Morere (New York: Springer), 107–126. doi: 10.1007/978-1-4614-5269-0_6

Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P., and Kroll, J. F. (2011). When deaf signers read English: do written words activate their sign translations? *Cognition* 118, 286–292. doi: 10.1016/j.cognition.2010.11.006

Murdoch, B. B. (1962). The serial position effect of free recall. *J. Exp. Psychol.* 64, 482–488. doi: 10.1037/h0045106

Nation, K., and Snowling, M. J. (1998a). Individual differences in contextual facilitation: evidence from dyslexia and poor reading comprehension. *Child Dev.* 69, 996–1011. doi: 10.1111/j.1467-8624.1998.tb06157.x

Nation, K., and Snowling, M. J. (1998b). Semantic processing and the development of word-recognition skills: evidence from children with reading comprehension difficulties. *J. Mem. Lang.* 39, 85–101. doi: 10.1006/jmla.1998.2564

Padden, C., and Ramsey, C. (2000). "American sign language and reading ability in deaf children," in *Language Acquisition by Eye*, eds C. Chamerlain, J. P. Morford, and R. I. Mayberry (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 165–189.

Parrila, R., Kirby, J. R., and McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: longitudinal predictors of early reading development? *Sci. Stud. Read.* 8, 3–26. doi: 10.1207/s1532799xssr0801_2

Pennington, B. F., and Bishop, D. V. (2009). Relations among speech, language, and reading disorders. *Annu. Rev. Psychol.* 60, 283–306. doi: 10.1146/annurev.psych.60.110707.163548

Perfetti, C. A., and Hart, L. (2001). "The lexical quality hypothesis," in *Precursors of Functional Literacy*, Vol. 11, eds C. Elbro, P. Reitsma, and L. Verhoeven (Amsterdam: John Benjamins), 67–86.

Perfetti, C. A., Landi, N., and Oakhill, J. (2005). "The acquisition of reading comprehension skill," in *The Science of Reading: A Handbook*, eds M. J. Snowling and C. Hulme (Oxford: Blackwell Publishing Ltd). 227–247.

Piñar, P., Dussias, P. E., and Morford, J. P. (2011). Deaf readers as bilinguals: an examination of deaf readers' print comprehension in light of current advances in bilingualism and second language processing. *Lang. Linguist. Compass* 5, 691–704. doi: 10.1111/j.1749-818X.2011.00307.x

Prinz, P. M., and Strong, M. (1998). ASL proficiency and English literacy within a bilingual deaf education model of instruction. *Top. Lang. Disord.* 18, 47. doi: 10.1097/00011363-199808000-00006

R Development Core Team. (2010). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rundus, D., and Atkinson, R. C. (1970). Rehearsal processes in free recall: a procedure for direct observation. *J. Verbal Learn. Verbal Behav.* 9, 99–105. doi: 10.1016/S0022-5371(70)80015-9

Scarborough, H. S. (2009). "Connecting early language and literacy to later reading (dis) abilities: evidence, theory, and practice," in *Approaching Difficulties in Literacy Development: Assessment, Pedagogy, and Programmes*, eds F. Fletcher-Campbell, J. Soler and G. Reid (Milton Keynes: Sage), 23–39.

Shankweiler, D., and Liberman, I. Y. (eds). (1989). "How problems of comprehension are related to difficulties in decoding," in *Phonology and Reading Disability: Solving the reading Puzzle*, (Ann Arbor: University of Michigan Press), 35–67 .

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biol. Psychiatry* 55, 926–933. doi: 10.1016/j.biopsych.2003.12.019

Siegel, L. S., and Linder, B. A. (1984). Short-term memory processes in children with reading and arithmetic learning disabilities. *Dev. Psychol.* 20, 200–207. doi: 10.1037/0012-1649.20.2.200

Sinatra, R. C., Stahl-Gemake, J., and Berg, D. N. (1984). Improving reading comprehension of disabled readers through semantic mapping. *Read. Teach.* 38, 22–29.

Snowling, M. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Adoles. Ment. Health* 3, 4–11. doi: 10.1111/1475-3588.00201

Stahl, S. A., and Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *J. Educ. Psychol.* 86, 221–234. doi: 10.1037/0022-0663.86.2.221

Sterne, A., and Goswami, U. (2000). Phonological awareness of syllables, rhymes, and phonemes in deaf children. *J. Child Psychol. Psychiatry Allied Discipl.* 41, 609–625. doi: 10.1111/1469-7610.00648

Supalla, T., Hauser, P. C., and Bavelier, D. (2014). Reproducing American Sign Language sentences: cognitive scaffolding in working memory. *Front. Psychol.* 5:859. doi: 10.3389/fpsyg.2014.00859

Swanson, H. L. (1999). Reading comprehension and working memory in learning-disabled readers: is the phonological loop more important than the executive system? *J. Exp. Child Psychol.* 72, 1–31. doi: 10.1006/jecp.1998.2477

Swanson, H. L., and Ashbaker, M. H. (2000). Working memory, short-term memory, speech rate, word recognition and reading comprehension in learning disabled readers: does the executive system have a role? *Intelligence* 28, 1–30. doi: 10.1016/S0160-2896(99)00025-2

Swanson, H. L., and Howell, M. (2001). Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. *J. Educ. Psychol.* 93, 720. doi: 10.1037/0022-0663.93.4.720

Traxler, C. B. (2000). The stanford achievement test 9th edition: national norming and performance standards for deaf and hard of hearing students. *J. Deaf Stud. Deaf Educ.* 5, 337–348. doi: 10.1093/deafed/5.4.337

Trezek, B. J., Wang, Y., and Paul, P. V. (2010). *Reading and Deafness: Theory, Research, and Practice*. Clifton Park, NY: Cengage Learning.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative procedural model. *Bilingualism* 4, 105–122. doi: 10.1017/S1366728901000220

Ullman, M. T. (2005). "A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model," in *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice*, ed. C. Sanz (Washington, DC: Georgetown University Press), 141–178.

Wagner, R. K., and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101, 192–212. doi: 10.1037/0033-2909.101.2.192

Wagner, R. K., Torgesen, J. K., and Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: new evidence of bidirectional causality from a latent variable longitudinal study. *Dev. Psychol.* 30, 73–87. doi: 10.1037/0012-1649.30.1.73

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Dev. Psychol.* 33, 468–478. doi: 10.1037/0012-1649.33.3.468

Walden, B. E., Grant, K. W., and Cord, M. T. (2001). Effects of amplification and speechreading on consonant recognition by persons with impaired hearing. *Ear Hear.* 22, 333–341. doi: 10.1097/00003446-200108000-00007

Waters, H. S., and Waters, E. (1976). Semantic processing in children's free recall: evidence for the importance of attentional factors and encoding variability. *J. Exp. Psychol.* 2, 370–380. doi: 10.1037/0278-7393.2.4.370

Yurkowski, P., and Ewoldt, C. (1986). A case for the semantic processing of the deaf reader. *Am. Ann. Deaf* 131, 243–247. doi: 10.1353/aad.2012.0795

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Hirshorn, Dye, Hauser, Supalla and Bavelier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 05 May 2015
doi: 10.3389/fpsyg.2015.00527

Deaf children's non-verbal working memory is impacted by their language experience

Chloë Marshall[1†], *Anna Jones*[2†], *Tanya Denmark*[2], *Kathryn Mason*[2], *Joanna Atkinson*[2], *Nicola Botting*[3] and *Gary Morgan*[3*]

[1] Department of Psychology and Human Development, UCL Institute of Education, University College London, London, UK, [2] Deafness, Cognition and Language Research Centre, University College London, London, UK, [3] Division of Language and Communication Sciences, City University London, London, UK

OPEN ACCESS

Edited by:
Mary Rudner, Linköping University, Sweden

Reviewed by:
Mireille Besson, Centre National de la Recherche Scientifique, France
Matt Hall, University of California, San Diego, USA

***Correspondence:**
Gary Morgan, Division of Language and Communication Sciences, City University London, Northampton Square, London EC1V 0HB, UK
g.morgan@city.ac.uk

[†] *These authors have contributed equally to this work.*

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *30 January 2015*
Accepted: *13 April 2015*
Published: *05 May 2015*

Citation:
Marshall C, Jones A, Denmark T, Mason K, Atkinson J, Botting N and Morgan G (2015) Deaf children's non-verbal working memory is impacted by their language experience. Front. Psychol. 6:527. doi: 10.3389/fpsyg.2015.00527

Several recent studies have suggested that deaf children perform more poorly on working memory tasks compared to hearing children, but these studies have not been able to determine whether this poorer performance arises directly from deafness itself or from deaf children's reduced language exposure. The issue remains unresolved because findings come mostly from (1) tasks that are verbal as opposed to non-verbal, and (2) involve deaf children who use spoken communication and therefore may have experienced impoverished input and delayed language acquisition. This is in contrast to deaf children who have been exposed to a sign language since birth from Deaf parents (and who therefore have native language-learning opportunities within a normal developmental timeframe for language acquisition). A more direct, and therefore stronger, test of the hypothesis that the type and quality of language exposure impact working memory is to use measures of non-verbal working memory (NVWM) and to compare hearing children with two groups of deaf signing children: those who have had native exposure to a sign language, and those who have experienced delayed acquisition and reduced quality of language input compared to their native-signing peers. In this study we investigated the relationship between NVWM and language in three groups aged 6–11 years: hearing children ($n = 28$), deaf children who were native users of British Sign Language (BSL; $n = 8$), and deaf children who used BSL but who were not native signers ($n = 19$). We administered a battery of non-verbal reasoning, NVWM, and language tasks. We examined whether the groups differed on NVWM scores, and whether scores on language tasks predicted scores on NVWM tasks. For the two executive-loaded NVWM tasks included in our battery, the non-native signers performed less accurately than the native signer and hearing groups (who did not differ from one another). Multiple regression analysis revealed that scores on the vocabulary measure predicted scores on those two executive-loaded NVWM tasks (with age and non-verbal reasoning partialled out). Our results suggest that whatever the language modality—spoken or signed—rich language experience from birth, and the good language skills that result from this early age of acquisition, play a critical role in the development of NVWM and in performance on NVWM tasks.

Keywords: deafness, language, British Sign Language, working memory

Introduction

Working memory is the capacity to encode, store, manipulate and recall information, and is essential for cognition (Baddeley and Hitch, 1974). As Hirshorn et al. (2012 p. 85) write, "One would be hard pressed to name any higher level cognitive ability that does not foundationally depend on holding information in memory and being able to manipulate and integrate it with knowledge from long-term memory." Not surprisingly, therefore, individual differences in working memory are associated with variation in such diverse activities as reasoning ability (Kyllonen and Christal, 1990), the acquisition of computer programming skills (Shute, 1991), and a whole set of activities that require language, such as reading comprehension (Daneman and Carpenter, 1980), novel word learning (Kwok and Ellis, 2014), syntactic processing (King and Just, 1991), second language learning (Kormos and Sáfár, 2008), acquiring an artificial language (Kapa and Colombo, 2014), and even adjusting to non-native speakers' lexical reference (Lev-Ari, 2015). Furthermore, individual differences in children's working memory are closely linked to their academic achievement (Alloway et al., 2005; Engel de Abreu et al., 2014). In the recent literature, the term working memory has been used to describe only the complex and executive-loaded elements of memory, i.e., where concurrent maintenance and processing of information are required for task completion. The focus of the study reported in the current paper is the nature of the association between language and working memory in the wider sense, although we were particularly interested in the complex and executive-loaded tasks.

As is often the case when trying to understand the nature of associative relationships between cognitive variables, it is far from straightforward to establish causal direction, i.e., whether differences in working memory drive individual differences in language during development, or vice versa. Longitudinal studies of children's vocabulary size have suggested that working memory ability does indeed drive vocabulary development rather than the other way round (Avons et al., 1998). Mechanistically, the claim is that the phonological loop (a component of phonological working memory; Baddeley and Hitch, 1974) provides a temporary means of storing new words, before they are consolidated in phonological long term memory (Baddeley et al., 1998). However, the strength of working memory as a predictor of vocabulary size declines with age (Gathercole et al., 1992) and is not found in all studies (Melby-Lervag et al., 2012).

A window onto the question of whether the causal influence might also operate in the opposite direction, i.e., whether individual differences in language can drive differences in working memory, comes from deaf children whose language learning experience is very different from that of the vast majority of children. The incidence of significant congenital deafness is about 1 in 1000 live births in most developed countries, including the UK, although it may be 3–4 times higher in certain communities or parts of the UK (Davis et al., 1997). Even mild deafness (defined as a hearing loss of 21–40 decibels) can cause difficulties accessing spoken language and have a detrimental effect on linguistic development. Hearing aids and cochlear implant technology, while improving rapidly, do not offer access to the same quality of speech that hearing children obtain naturally (Faulkner and Pisoni, 2013).

Sign languages such as British Sign Language (BSL) *do* offer a fully accessible language form to deaf children who do not have co-occurring visual impairment, but the vast majority of deaf children (over 90%; Lederberg and Mobley, 1990) are born to hearing non-signing parents. This means that even in cases where hearing parents learn BSL and sign with their children from an early age, the quality and quantity of language input and interaction that they are able to provide is likely to be impoverished compared to that provided by deaf signing parents. Nevertheless, for deaf children born to deaf signing parents, who receive sign language input from birth, language acquisition can show remarkable parallels in onset, rate and patterns of development compared to hearing children who are learning spoken languages (see Chamberlain et al., 2000; Morgan and Woll, 2002; Schick et al., 2005 for reviews). Deaf children of deaf parents (i.e., native signers) are therefore a very interesting population theoretically, but they are also very difficult to recruit to research studies. Not only are there a very small number of children in this group, but measuring their skills requires carefully-designed tasks and a researcher fluent in the particular sign language under consideration (Lieberman and Mayberry, 2015).

The diversity of language input in the deaf population, both with respect to age of access to language (from birth, later in infancy/childhood) and language form (signed or spoken), allows researchers to investigate how individual differences in linguistic input can impact on working memory development. In the remainder of this introduction, we review studies that have investigated working memory in deaf adults and children, identify the gaps in that literature, and motivate our own study.

A theme in the research literature on deaf people's working memory to date is a division between two types of studies: those that have investigated memory for spoken material, and those that have studied memory for signed and/or non-linguistic visuo-spatial material. Measurement of working memory across modalities requires serious consideration. It cannot be assumed that performance on a task presented in two different modalities is directly comparable. Likewise it cannot be assumed that two tasks presented in the same modality are directly comparable. As we discuss below, both modality and the nature of the material affect recall in working memory tasks.

It is perhaps not surprising that studies where material is presented auditorally find poorer recall by deaf participants in comparison to hearing participants. For example, Fagan et al. (2007) studied deaf children aged 6–14 years who received a cochlear implant between the ages of 1 and 6 years. Group means on spoken forward and backward digit span tasks were significantly lower than the standardized mean, with half the sample scoring below 1 SD from the mean on the forward task and the majority scoring below 1 SD from the mean on the backward task. Furthermore, scores on both span tasks were moderately correlated with vocabulary comprehension and non-word/rare-word reading scores. In another study, Burkholder and Pisoni (2003) divided deaf cochlear-implant users (aged 8–9 years) into two groups, according to whether they used just oral language or whether they used total communication (i.e., using manual sign

and lip reading strategies, in addition to speech), and compared them to a group of hearing children on spoken digit span tasks. Both deaf groups performed significantly more poorly than the hearing group. The digit span disadvantage for deaf participants has been found even when the task bypasses listening/speaking by being presented in written form (Parasnis et al., 1996), and when letters are used instead of digits (Wallace and Corballis, 1973).

A disadvantage for serially-presented linguistic material is also found when deaf participants undertake the digit span or letter span task in a sign language. Deaf native American Sign Language (ASL) signers recall on average only 5 ± 1 digits in forward tasks, compared to hearers who recall an average of 7 ± 2 digits (Boutla et al., 2004; Bavelier et al., 2006). Hall and Bavelier (2010, p. 54) have concluded that "speech-based representations are better suited for the specific task of perception and memory encoding of a series of unrelated verbal items in serial order through the phonological loop." Conway et al. (2009) go further and propose the "auditory scaffolding hypothesis," whereby one's experience with sound helps provide a scaffold for the development of those general cognitive abilities that are required for the representation of temporal or sequential patterns. However, Bavelier and colleagues' work shows that *hearing* English-ASL bilingual adults also show the same disadvantage for sign span compared to spoken span (Bavelier et al., 2008), which challenges the auditory scaffolding hypothesis because these individuals have had rich auditory input since birth. In any case, it is clear that performance on spoken serial recall tasks may not be directly comparable to performance on signed serial recall tasks.

For non-linguistic material that is *not* processed using the phonological loop, but which, like linguistic material, is serial in nature, deaf signers have been shown to have an advantage compared to other groups. Deaf adult signers have longer forward spans than hearing non-signers on the visuo-spatial Corsi Block Test (Geraci et al., 2008). Wilson et al. (1997) showed that the advantage for deaf signers over hearing non-signers in the Corsi Block Test was also evident in 8–10 year-old children. Evidence that the working memory advantage might arise from using sign language, rather than from being deaf, comes from studies by Capirci et al. (1998) and Parasnis et al. (1996). The former study demonstrated that hearing children who were taught sign language at school performed on non-verbal working memory tasks better after 1 year than hearing children who were taught a spoken language (Capirci et al., 1998), while the latter study found that deaf orally-educated children did not have an advantage over hearing children (Parasnis et al., 1996).

When serial recall of material is not the only requirement of the working memory task, or indeed is not required at all, then the pattern of results looks different again. Differences have *not* been found between deaf signers and hearing non-signers on complex span tasks, which rely on some sort of processing of material in addition to serial maintenance. However, the difficulty of complex span tasks means that to date in the deafness and sign language literature they appear to have only been carried out with adults (e.g., Boutla et al., 2004; Andin et al., 2013).

In summary, several recent studies have suggested that deaf children perform more poorly on working memory tasks compared to hearing children, but they have not been able to determine whether this poorer performance arises directly from deafness itself or from deaf children's reduced language exposure. The underlying cause of deaf children's poor task performance remains unresolved because findings come mostly from (1) tasks that are verbal as opposed to non-verbal (e.g., Burkholder and Pisoni, 2003; Fagan et al., 2007) and (2) deaf children who use spoken communication and who may therefore have experienced impoverished language input or have language development delay (e.g., Burkholder and Pisoni, 2003; Fagan et al., 2007; Figueras et al., 2008; Beer et al., 2011; Hintermair, 2013). Such a group may potentially perform differently on working memory tasks compared to deaf children who have been exposed to a sign language since birth from Deaf parents (and who therefore have native language-learning opportunities within a normal developmental timeframe for language acquisition). The role of age of language exposure in the wider neuro-cognitive abilities of deaf individuals has also been highlighted (Campbell et al., 2014). Moreover, studies using complex span tasks have not been reported, to the best of our knowledge, with deaf children. As mentioned earlier, recruiting and testing deaf children with a range of language experiences, and particularly those who are native signers, is a challenging task. However, doing so provides important contrasts which enable us to start unpacking the influences of auditory experience and language background.

A more direct, and consequently stronger, test of the relationship between type and quality of language exposure and working memory is therefore to use measures of non-verbal working memory and to compare hearing children with two groups of deaf signing children: those who have had native exposure to a sign language, and those who have experienced delayed acquisition and reduced quality of sign language input compared to their native-signing peers. This is exactly what we set out to do in the present study. If it is language experience rather than deafness that impacts on working memory, then native deaf signers should pattern like hearing children and both groups should perform better than non-native signers. Furthermore, scores on language tasks should correlate with working memory scores. If, however, it is lack of auditory experience that causes poor working memory, or if it is the case that comorbid memory difficulties occur with deafness, then both deaf groups should perform worse than the hearing group. If neither language experience nor deafness has an impact on working memory, then the three groups would not be expected to differ from one another, and no relationship should be found between language and working memory scores.

Methods

Participants

Twenty seven deaf children aged 6–11 years old (16 boys) were recruited. All had profound and/or severe hearing loss in both ears, with the majority ($n = 24$) being profoundly deaf in both ears. All had been born deaf (i.e., none had been deafened in early childhood by meningitis, for example, and therefore none had early access to auditory input). However, none had additional learning difficulties, according to teacher and/or parental report.

All 27 deaf participants used BSL regularly, but had different levels of exposure to BSL and different degrees of BSL use. Based

on their exposure to, and use of, BSL, they were divided into two groups: native signers ($n = 8$) and non-native signers ($n = 19$). To be included in the native signer group, participants had to have at least one deaf parent (some also had one or more deaf siblings, but this was not a requirement for inclusion) and to have been exposed to BSL from their parent(s) since birth. In addition, the parents of these children had to report that BSL was the language in which their child preferred to communicate and was the language in which the child communicated with his/her deaf parent(s). Although not part of the selection criteria, the eight children in this group (5 boys) were all reported to mix regularly with deaf adults and either half or the majority of their friends were reported to be deaf. Please see **Table 1** for further details.

The remaining 19 deaf participants (11 boys) were considered to be non-native signers. This group was characterized by a later age of acquisition of BSL than the native-signer group ($M = 2;11$ years, $SD = 2;2$ years, range = 0;7–9;0 years), and the majority ($n = 13$) were reported to use sign-supported English (SSE) or spoken English alongside BSL as their preferred language and with their hearing parents. As **Table 2** shows, this was a more heterogeneous group with respect to language background and current language use than the native-signer group, as is to be expected.

Twenty eight hearing participants of the same age—6 to 11 years (16 boys)—were also recruited. All were reported by parents/and or teachers to have no hearing difficulties or learning difficulties of any kind, and all had English as their first language.

The mean age of the deaf participants was 9;2 years ($SD = 1;8$), and of the hearing participants was 9;0 ($SD = 1;5$). There was no significant age difference between the deaf and hearing groups, $t_{(53)} = 0.320$, $p = 0.751$. However, the native signers ($M = 8;0$, $SD = 0;11$) were significantly younger than the non-native signers ($M = 9;7$, $SD = 1;9$), $t_{(25)} = 2.391$, $p = 0.025$, and marginally younger than the hearing participants, $t_{(34)} = 1.802$, $p = 0.080$. The non-native signer and hearing groups did not differ on age, $t_{(45)} = 1.289$, $p = 0.204$.

With respect to non-verbal reasoning, as measured by the matrix reasoning subset of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), both groups had mean T-scores in the normal range (mean = 50, $SD = 10$), $M_{\text{deaf}} = 52.33$ ($SD_{\text{deaf}} = 10.57$) and $M_{\text{hearing}} = 55.79$ ($SD_{\text{hearing}} = 8.48$), and the scores did not differ significantly from one another, $t_{(53)} = 1.338$, $p = 0.186$. Within the deaf group, the native signer subgroup ($M = 62.25$, $SD = 7.01$) had significantly higher T-scores than the non-native signers ($M = 48.16$, $SD = 8.96$), $t_{(25)} = 3.954$, $p = 0.001$, and marginally higher T-scores than the hearing group, $t_{(34)} = 1.967$, $p = 0.057$. The hearing group had higher T-scores than the non-native signers, $t_{(45)} = 2.959$, $p = 0.005$.

Materials

Working Memory Tasks

Two working memory tasks, namely the Spatial Span Task (Wechsler and Naglieri, 2006) and the Odd One Out Span Task (Henry, 2001), were selected after piloting as they require a minimal amount of verbal instruction and only non-verbal responses (i.e., pointing).

The *Spatial Span Task* (from the Wechsler Nonverbal Scale of Ability, Wechsler and Naglieri, 2006) is a measure of visuo-spatial short-term working memory similar to the Corsi Block Test. A set of nine identical blue blocks is affixed to a white board in an unstructured array. The examiner can view a number on each of the blocks and is seated directly opposite to the child being tested. Children are instructed to tap a sequence of blocks in the same order as the examiner in the "forward" test, and in the reverse order in the "backward" test. Children are administered two trials for each sequence length, beginning with two blocks, ranging up

TABLE 1 | Language background of deaf native signers.

Code*	Age	Deaf family members	Parent-child language preference	Child's language preference	Age of acquisition of BSL	Who did child learn BSL from?	Are child's friends deaf?	Does child mix with deaf adults?	Level of deafness in left ear	Level of deafness in right ear	Does child wear hearing aids?	Does child have a cochlear implant?
B01	7;1	Parent(s)	BSL	BSL	From birth	Deaf parent(s)	Most deaf	Yes	Profound	Profound	Sometimes	No
G02	8;0	Parent(s)	BSL	BSL	From birth	Deaf parent(s)	Most deaf	Yes	Profound	Profound	Sometimes	No
G03	7;1	Parent(s) + sibling(s)	BSL	BSL	From birth	Deaf parent(s)	Equal deaf/hearing	Yes	Profound	Profound	Sometimes	No
B04	7;6	Parent(s) + sibling(s)	BSL	BSL	From birth	Deaf parent(s)	Most deaf	Yes	Profound	Profound	Sometimes	No
B05	8;1	Parent(s) + sibling(s)	BSL	BSL	From birth	Deaf parent(s)	Equal deaf/hearing	Yes	Profound	Profound	No	No
G06	8;11	Parent(s) + sibling(s)	BSL	BSL	From birth	Deaf parent(s)	Most deaf	Yes	Profound	Profound	Yes	No
B07	7;10	Parent(s) + sibling(s)	BSL	BSL	From birth	Deaf parent(s)	Equal deaf/hearing	Yes	Severe	Severe	Yes	No
B08	9;9	Parent(s) + sibling(s)	BSL	BSL	From birth	Deaf parent(s)	Most deaf	Yes	Profound	Profound	Sometimes	No

B, Boy and G, Girl.

TABLE 2 | Language background of deaf non-native signers.

Code*	Age	Deaf family members	Parent-child language preference	Child's language preference	Age of acquisition of BSL (years)	Who did child learn BSL from?	Are child's friends deaf?	Does child mix with deaf adults?	Level of deafness in left ear	Level of deafness in right ear	Does child wear hearing aids?	Does child have a cochlear implant?
G09	11;9	Older sibling	BSL	BSL	5	Deaf older sibling and school	Most deaf	Yes	Profound	Profound	Sometimes	No
G10	11;9	Twin sibling	BSL + spoken English	BSL + spoken English	4	School and speech therapist	Most deaf	Yes	Profound	Profound	No	Yes
G11	7;3	Other(s)	BSL	BSL + spoken English	0.8	Mother	Equal deaf/hearing	Yes	Profound	Profound	No	Yes
B12	10;4	None	BSL	BSL	0.6	Unknown	Equal deaf/hearing	Yes	Profound	Profound	Sometimes	Yes
B13	9;10	None	BSL	BSL	4	School	Equal deaf/hearing	No	Profound	Profound	No	No
B14	7;1	None	BSL	BSL	1	Parents and language aide	Most hearing	Yes	Profound	Profound	Not any more	No
B15	6;8	None	BSL	BSL	3	Mother	Equal deaf/hearing	Yes	Profound	Profound	No	Yes
B16	9;7	None	BSL	BSL	1	Parents; deaf and hearing adults	Most deaf	Yes	Profound	Profound	No	No
G17	9;4	None	SSE	SSE + BSL	2	School	Most deaf	No	Profound	Profound	Yes	No
G18	8;3	None	SSE	SSE + BSL	3	School	Most deaf	Yes	Severe	Severe	Yes	No
B19	8;6	None	SSE	SSE + BSL	6	School	Equal deaf/hearing	Yes	Profound	Profound	Yes	No
G20	10;7	None	BSL + spoken English	BSL	1.6	Mother and school	Equal deaf/hearing	Yes	Profound	Profound	No	Yes
B21	10;2	None	BSL + spoken English	BSL	2	Mother	Most deaf	Yes	Profound	Profound	Sometimes	No
B22	11;0	None	BSL + spoken English	SSE	3	Peripatetic teacher of the deaf	Most hearing	Yes	Profound	Profound	Yes	No
G23	9;5	None	BSL + spoken English	SSE + BSL	1	Parents and nursery	Most hearing	Yes	Profound	Profound	No	Yes
B24	11;10	None	BSL + spoken English	BSL + spoken English	4	School	Most deaf	No	Profound	Profound	Not any more	Yes
B25	11;8	None	spoken English	BSL	9	School	Most deaf	No	profound	profound	yes	yes
B26	6;9	None	spoken English	BSL + spoken English	3	School	Most deaf	Yes	severe	severe	yes	no
G27	10;11	None	spoken English	BSL + spoken English	1	Parents and school	Most deaf	Yes	profound	profound	no	yes

**B, Boy; G, Girl.*

to a span of nine. Two trials of each sequence length are administered, and the test is terminated once both trials of the same sequence length are failed. The task begins with two practice trials in both the spatial span forward and backward conditions to ensure that the child understands the task. One point is awarded for each sequence accurately repeated.

The *Odd One Out Span Task* (Henry, 2001) is a measure of executive-loaded visuo-spatial working memory. It is presented

in PowerPoint and comprises 63 slides, each displaying a set of three shapes. On each of the slides, two of the shapes are identical, and one is slightly different: the "odd one out." The examiner shows the child a slide and asks them to identify which shape is the odd one out. The child is instructed to try to remember the location of this shape. The following slide contains an empty grid with three boxes, and the child is asked to point to the empty box in the same location as the shape that they have just seen. After four single-item trials have been displayed, the child is shown two sets of shapes in a row. There then follows a slide with two empty grids, one on top of the other. The child is instructed to point to the empty boxes in the same location as the two "odd" shapes they have previously seen, in the same order that they were presented. If the child initially verbalizes or signs their answer (e.g., *left*, *middle*, etc.), they are reminded that they need to point to the location of the shape. Trial length increases sequentially in blocks of four with a maximum of six sets of shapes. Once the child makes two errors within a block, the test is terminated. The total number of trials correctly recalled is then calculated. Before the test begins, two practice trials are administered to illustrate the task procedure: a single-item and a two-item trial. Correct responses to the practice items are indicated to the child if they do not initially answer correctly.

Language Tasks

We used three tests of language, of which two were new adaptations of existing measures. An adapted version of the *Expressive One Word Picture Vocabulary Test* (EOWPVT; Brownell, 2000) was used to test single word vocabulary production. The full test was initially administered as per the instruction manual. The children are presented with single pictures that test knowledge of primarily simple nouns (e.g., *train*, *pineapple*, *kayak*), but also some verbs (e.g., *eating*, *hurdling*), and category labels (e.g., *fruit*, *food*). After four practice items, the test begins at various starting points depending on the child's age. Eight items must be labeled correctly in succession, and the experimenter works backwards if necessary until the basal is achieved. The test finishes when the child gets six successive incorrect answers. The EWOPVT was developed in the USA and so a few pictures ($n = 3$) were substituted with alternative pictures to make the test more culturally relevant for children in the UK (e.g., *raccoon* → *badger*). In order that the EOWPVT could be used to assess the vocabulary of both hearing and deaf children who communicate in BSL, it was necessary to exclude a number of test items that do not exist in BSL (e.g., *cactus*, *banjo*, "*musical instruments*" as a collective term). This list of 15 excluded items was established by administering the test to three native signing Deaf adults who primarily communicated in BSL. These items were then deducted from the children's total raw scores.

The *BSL Narrative Production Test* (Herman et al., 2004) was designed to assess deaf children's (age 4–11 years) expressive language by eliciting a narrative in BSL. The child first watches a short, silent video (on a DVD) acted out by two deaf children. Participants are instructed to watch it carefully as they are going to be asked to tell the story once the video has finished. The experimenter leaves the room while the child watches the video and returns once it has finished. The experimenter asks the child to tell the story. The aim is to elicit a spontaneous story, so no further prompting is given other than asking, "is there anything else?" to check that the child has finished. The child's narrative is videotaped for subsequent scoring. The test is scored based upon three components: (1) the content of the story (i.e., the level of detailed information included in their narrative); (2) story structure (i.e., introducing the participants and setting the scene, reporting the key events leading to the climax of the story, and detailing the resolution of the story at the end); (3) aspects of BSL grammar (including use of spatial location, person and object classifiers and role shift). The narratives were scored by an experimenter who was fluent in BSL and had completed the training course required for administrators/coders of the test.

Hearing control group children were also tested on their narrative skills using the same video to elicit a spontaneous story in spoken English. As the original story is told only through gesture and action, this prompted the hearing children to use some gesture in their story retellings e.g., when describing the boy demanding food from the girl, a child may say: "Then he went like that [gestures putting out hand]." These gestures were included in the scoring of the story content. Because English and BSL grammar systems are very different, only narrative content and structure were scored for the purpose of this study. The reliability of the use of the test in spoken English was investigated with composite scores of structure and content. Twenty-four of the narratives were scored by two trained testers, showing good inter-rater reliability ($r = 0.97$, $p < 0.001$). Ten of the narratives were scored a second time by the same scorer, showing high intra-rater reliability ($r = 0.98$, $p < 0.001$). The internal consistency between the content and structure items of the measure was also high ($r = 0.90$, $p < 0.001$).

The *Language Proficiency Profile-2* (LPP-2; Bebko and McKinnon, 1993) is a questionnaire completed by a person who is familiar with the child's language skills. The aim is to provide an overall evaluation of linguistic and communicative skills of deaf children, regardless of the specific language or modality in which they communicate (i.e., BSL, signed supported English, spoken English, etc.). Most usually the parents, but occasionally the teacher ($n = 3$, all in the deaf group), of the children participating in this study completed this questionnaire. The LPP-2 comprises five categories: (1) Form: structure of the language e.g., single words/signs in the early stages, later developing into the ability to produce short narratives; (2) Use: functions of language i.e., to interact or gain the attention of others etc.; (3) Content: the type of objects, actions and relationships that exist in the child's communication e.g., referring to the existence/disappearance of objects, information about denial or causality etc.; (4) Reference: the ability of the child to describe events beyond the present context; and (5) Cohesion: how effectively the child adapts their communication to the listener e.g., modification of syntax to account for the perspective, knowledge and opinion of their conversational partner (Bebko and McKinnon, 1993). Each item is rated on a scale with five options: past this level, yes, emerging, not yet, or unsure. Up to 18 points are available for form, 24 for Content, 22 for Reference, 22 for Cohesion, and 26 for Use. We combined the scores on the five sections to give an aggregate score (out of a possible 112 points). The LPP-2 takes approximately 15 min to

complete and has been shown to have good concurrent validity with language measures used with both deaf and hearing children (Bebko et al., 2003).

Non-verbal Reasoning Task

Finally, the *Matrix Reasoning* subtest of the *Wechsler Abbreviated Scale of Intelligence* (WASI; Wechsler, 1999) was also administered as a control measure. Matrix reasoning is a performance IQ assessment of non-verbal fluid ability. The child is presented with a pattern with a missing section and is instructed to select the correct response from five potential choices. The starting and stopping points are determined by the participant's age, and the matrices become increasingly difficult to solve. The test begins with two practice items to ensure the child has understood the task. The test is terminated when four successive answers, or four out of five successive answers, are incorrect.

Procedure

Prior to data collection, written parental consent was obtained and the LPP-2 questionnaire was also completed by parents. (For three children, all in the deaf group, the LPP-2 was completed instead by the child's teachers). Face-to-face consent was obtained from the children at the start of the testing session. The children were tested individually in a quiet room, either at school or at home, in a session lasting 35–45 min. (The approximate timings for each test were: Corsi blocks—5 min; Odd one out task—5 to 10 min; Narrative—5 to 10 min; Vocabulary—10 min; WASI matrix reasoning—10 min). The entire session was videotaped. Testing of the deaf children was carried out by an adult hearing native user of BSL, who is highly experienced in communicating with deaf children. The hearing children were also tested by this adult and by three additional trained hearing experimenters. Standardized test instructions (translated into BSL) were used for all of the tests. As mentioned earlier, the tasks required minimal verbal/signed instruction, and sufficient practice trials were included to ensure understanding of the task requirements. It was ensured that lighting conditions were good and that children could see the experimenter clearly to view lip movements. The tests were administered in the same order for all participants to ensure that possible test-order effects would be consistent across groups.

Results

Data were missing from one hearing child for the BSL Narrative Production test, and from five deaf and two hearing children for LPP-2. Otherwise the dataset was complete. We present three sets of analyses. First, we compare the entire group of deaf children to the group of hearing children on all language and working memory measures. Secondly, we split the deaf group according to language experience into native and non-native signer groups, and compare them to the hearing children on all language and working memory measures. Finally, we investigate whether language scores predict working memory scores in the deaf and hearing children considered together.

Comparison of Deaf vs. Hearing Groups

Raw scores for the deaf and hearing groups on the language and working memory tasks are presented in **Table 3**. A series of independent samples *t*-tests was carried out to test for group differences. Because the groups did not differ for age and WASI matrix reasoning score, those factors were not controlled for in this analysis.

For the working memory tasks, the hearing group significantly outscored the deaf group on two measures: Spatial Span Backward, $t_{(53)} = 2.345$, $p = 0.023$, and Odd One Out, $t_{(53)} = 2.650$, $p = 0.011$. There were no group differences on the Spatial Span Forward task, $t_{(53)} = 1.231$, $p = 0.224$.

For the language tasks, the hearing group significantly outscored the deaf group on two measures: the Expressive One Word Picture Vocabulary Test, $t_{(53)} = 6.883$, $p < 0.001$, and the Language Proficiency Profile, $t_{(46)} = 3.401$, $p = 0.001$. There were, however, no group differences for BSL Narrative: Content, $t_{(52)} = 0.803$, $p = 0.426$, and BSL Narrative: Structure, $t_{(52)} = 0.193$, $p = 0.849$. Overall, therefore, where group differences were found on language and working memory measures, they favored the hearing group.

Comparison of Native Signer and Non-Native Signer vs. Hearing Groups

Table 4 presents the results of the language and working memory tasks for the three groups separately. Because the non-native signers were significantly older than the native signers, and

TABLE 3 | Mean (standard deviation) raw scores for the language and working memory measures, for the deaf and hearing groups.

	Measure (highest possible score)	Deaf children (*N* = 27)		Hearing children (*N* = 28)	
		Mean	(*SD*)	Mean	(*SD*)
Working memory	Spatial span: forwards (16)	6.22	(1.55)	6.75	(1.62)
	Spatial span: backwards (16)	4.93*	(1.98)	6.18	(1.98)
	Odd one out span (24)	8.85*	(4.03)	12.29	(5.45)
Language	Expressive one word picture vocabulary test (155)	64.44***	(15.76)	92.14	(14.06)
	BSL narrative assessment: content (16)	11.22	(3.06)	10.52	(3.38)
	BSL narrative assessment: structure (12)	9.37	(2.02)	9.26	(2.23)
	Language proficiency profile-2 (112)	97.27**	(15.86)	108.58	(5.52)

*The deaf group scores significantly lower than the hearing group: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.*

TABLE 4 | Estimated marginal means (standard error) for the language and working memory measures (controlling for age and WASI T-score), for the deaf native signer, deaf non-native signer, and hearing groups.

	Measure	Deaf native signers (*N* = 8)		Deaf non-native signers (*N* = 19)		Hearing children (*N* = 28)	
		Mean	(*SE*)	Mean	(*SE*)	Mean	(*SE*)
Working memory	Spatial span: forwards	7.00	(0.55)	5.82	(0.36)	6.80	(0.27)
	Spatial span: backwards	5.79	(0.70)	4.60*	(0.45)	6.16	(0.35)
	Odd one out span	10.43	(1.57)	8.40*	(1.03)	12.14	(0.78)
Language	Expressive one Word picture Vocabulary test	76.00**	(4.69)	59.90***	(3.06)	91.93	(2.32)
	BSL narrative assessment: content	11.45	(1.17)	11.23	(0.76)	10.44	(0.59)
	BSL narrative assessment: structure	9.73	(0.78)	9.28	(0.52)	9.22	(0.40)
	Language proficiency profile-2	110.46	(4.81)	91.93***	(2.93)	108.82	(2.11)

*Group scores significantly lower than the hearing group: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.*

because the non-native signers scored lower on the WASI matrix reasoning than both the native signers and the hearing group, we investigated group differences using ANCOVAs wherein we controlled for age and WASI score. Within each ANCOVA we carried out *post hoc* comparisons for group, using the Sidak correction to adjust for multiple comparisons.

For the working memory tasks there was a significant effect of group for the Spatial Span Backward task, $F_{(2, 54)} = 3.449$, $p = 0.040$. *Post hoc* tests revealed just one significant group difference: the non-native signers scored significantly more poorly than the hearing group, $p = 0.034$. Likewise, for the Odd One Out task, there was a significant effect of group, $F_{(2, 54)} = 4.187$, $p = 0.021$, with the only group difference being between the non-native signers and the hearing group, $p = 0.020$. There were no group differences for the Spatial Span Forward task, $F_{(2, 54)} = 2.474$, $p = 0.094$.

For the language tasks, the Expressive One Word Picture Vocabulary Test demonstrated a highly significant effect of group, $F_{(2, 54)} = 34.829$, $p < 0.001$, and this was driven by all three groups being significantly different from one another: non-native signer < native signer, $p = 0.029$, non-native signer < hearing, $p = 0.001$, and native signer < hearing, $p = 0.009$. For the Language Proficiency Profile, there was also a significant effect of group, $F_{(2, 47)} = 10.688$, $p = 0.001$, with the non-native signer group scoring significantly lower than both the native signer group, $p = 0.011$, and the hearing group, $p < 0.001$. There were no significant differences in scores between the native signer and hearing groups. As before, there were no group differences for scores on the BSL Narrative Test: Content, $F_{(2, 53)} = 0.542$, $p = 0.585$, and BSL Narrative Test: Structure, $F_{(2, 53)} = 0.174$, $p = 0.841$.

Using Language Scores to Predict Working Memory Scores

To explore the contribution of language to working memory test scores, a set of multiple regression analyses was carried out across all participants to predict scores on each working memory task. Expressive One Word Picture Vocabulary Test and Language Proficiency Profile scores were used as predictors. (The scores for content and structure in the BSL Narrative Production Test were not used because they had shown no group differences.) Age and WASI matrix reasoning T-scores were used as control predictors, and were entered into the model in a first step.

For the Spatial Span Backward task, a model with just age and WASI score entered as predictors of span scores was significant, $F_{(2, 47)} = 4.167$, $p = 0.022$, adjusted $R^2 = 0.119$. Both age ($\beta = 0.397$, $t = 2.680$, $p = 0.010$) and WASI score ($\beta = 0.299$, $t = 2.016$, $p = 0.050$) were unique predictors. When the two language measures were added, the model was a better fit, $F_{(4, 47)} = 4.465$, $p = 0.004$, adjusted $R^2 = 0.228$. In this new model, both age ($\beta = 0.273$, $t = 1.772$, $p = 0.083$) and WASI score ($\beta = 0.174$, $t = 1.196$, $p = 0.238$) lost their unique predictive power. Expressive One Word Picture Vocabulary score was a significant unique predictor ($\beta = 0.379$, $t = 2.208$, $p = 0.033$), but the Language Proficiency Profile score was not ($\beta = 0.029$, $t = 0.176$, $p = 0.861$).

The Odd One Out task showed a similar pattern to the Spatial Span Backward task. Age and WASI score entered together into the model were significant predictors of span scores, $F_{(2, 47)} = 8.192$, $p = 0.001$, adjusted $R^2 = 0.234$. Both age ($\beta = 0.497$, $t = 3.596$, $p = 0.001$) and WASI score ($\beta = 0.427$, $t = 3.092$, $p = 0.003$) were unique predictors. When the two language measures were added, the model showed an excellent fit, $F_{(4, 47)} = 11.368$, $p < 0.001$, and explained almost half of the variance in complex span (adjusted $R^2 = 0.469$). Age ($\beta = 0.329$, $t = 2.576$, $p = 0.014$) and WASI score ($\beta = 0.260$, $t = 2.149$, $p = 0.037$) remained significant predictors. Expressive One Word Picture Vocabulary score was also a significant unique predictor ($\beta = 0.510$, $t = 3.588$, $p = 0.001$), but the Language Proficiency Profile score was not ($\beta = 0.035$, $t = 0.261$, $p = 0.795$).

Finally, for the Spatial Span Forward task, age and WASI score entered together into the model were significant predictors of span scores, $F_{(2, 47)} = 6.456$, $p = 0.003$, adjusted $R^2 = 0.188$. Only age was a unique predictor ($\beta = 0.495$, $t = 3.480$, $p = 0.001$). WASI score was not a unique predictor, ($\beta = 0.071$, $t = 0.501$, $p = 0.619$). Adding the two language measures improved the model's fit, $F_{(4, 47)} = 5.746$, $p = 0.001$, adjusted $R^2 = 0.288$. Age remained a significant unique predictor ($\beta = 0.396$, $t = 2.677$, $p = 0.010$), but WASI score again was not ($\beta = -0.041$, $t = -0.294$, $p = 0.770$). Neither Expressive

One Word Picture Vocabulary score ($\beta = 0.318$, $t = 1.933$, $p = 0.060$) nor Language Proficiency Profile score ($\beta = 0.086$, $t = 0.553$, $p = 0.583$) were unique predictors of Spatial Span Forward score.

Summary of Results

For the Expressive One Word Picture Vocabulary Test we found significant group differences: the deaf group as a whole, and the native signer and non-native signer groups separately, scored more poorly than the hearing group. Furthermore, the non-native signers scored significantly lower than the native signers. For the Language Proficiency Profile, the pattern was a little different: the group difference between the deaf and hearing groups appeared to be driven by the poor performance of the non-native signer group.

The two executive-loaded working memory measures, the Spatial Span Backward and the Odd One Out Task, patterned like the Language Proficiency Profile: the group of deaf children as a whole and the non-native signer group separately scored lower than the hearing children, but the native signer group did not. For these two working memory measures, vocabulary as measured by the Expressive One Word Picture Vocabulary Test was a significant predictor of scores, beyond age and WASI matrix reasoning score, revealing an association between language and executive-loaded working memory.

However, for the BSL Narrative Production Test, scores for Narrative and Content revealed no group differences—the group of deaf children as a whole, and the two separate groups of deaf native and deaf non-native signers, scored at the same level as the hearing group. Similarly, for one of the working memory measures, the Spatial Span Forward task, we found no group differences. Thus, it is not inevitable that language and working memory performance is worse in deaf children compared to hearing—it depends on the nature of the task.

Discussion

In this study, we investigated the relationship between language and working memory by comparing three groups of children with different language experiences: hearing children, deaf native signers, and deaf non-native signers. These three groups allowed us to tease apart the impact of the quality of auditory experience vs. the impact of reduced language experience on working memory. If disturbances to auditory experience cause poor working memory, then both deaf groups should have performed worse than the hearing group. If it is language experience rather than deafness that impacts on working memory, then deaf native signers should have patterned like hearing children and both groups should have performed better than non-native signers. Furthermore, scores on language tasks should have correlated with working memory scores.

Our findings are consistent with this latter hypothesis that language experience, but not deafness *per se*, impacts on non-verbal, executive-loaded working memory. Although our group of deaf children as a whole performed more poorly than an age-matched group of hearing children on the two tasks that involved executive-loaded working memory (the Spatial Span Backward and the Odd One Out Tasks), this poor performance was driven by those deaf children who had had delayed and reduced language exposure by not having received signed language input from birth. The small subset of children ($n = 8$) who had learnt BSL under "native" language-learning conditions, i.e., from deaf signing parents, and who had rich language interactions throughout their childhood with family members, friends and at school, did not differ from the hearing group in their working memory scores. We do of course need to be cautious in our interpretation: our group of deaf native signers was small, as this is a rare population. Indeed, small sample sizes are prevalent in experimental studies of native signers' working memory (e.g. $n = 6$ in Wang and Napier, 2013; $n = 8$ in Krakow and Hanson, 1985; $n = 11$ in Wilson and Emmorey, 2006). Finally in our study, vocabulary, as indexed by the Expressive One Word Picture Vocabulary Test, was a strong predictor of scores on both executive-loaded working memory tasks when all children were considered together.

As discussed in the introduction, teasing apart causal relations between two variables over developmental time is not straightforward. For example, working memory has been extensively investigated in children with Specific Language Impairment (SLI). It has been argued that poor language directly impacts on working memory in children with SLI (van der Lely and Howard, 1993), but others have argued otherwise. In a more recent study Henry et al. (2012) administered the same non-verbal Odd One Out task as we used in our study (Henry, 2001) and a verbal working memory task, Listening Recall (Working Memory Test Battery for Children, Pickering and Gathercole, 2001). Henry et al. (2012) found that groups of children with poor language [both normal IQ (i.e., SLI) and low IQ] scored lower than typically developing children on both the verbal and non-verbal working memory tasks. In particular, performance on these tasks remained lower for the SLI group even when verbal IQ was entered in the regression analyses. Henry et al. (2012) conclude that their results are consistent with SLI being caused by a domain-general impairment rather than by an impairment specific to language (see also Ullman and Pierpont, 2005). However, these issues are difficult to tease apart in a population that is heterogeneous with respect to the severity and profile of language difficulties, and where it is possible that deficits of both domain-general and domain-specific (i.e., language) origin co-occur in the same child.

In contrast, the current study involved groups of children where language experience is affected in two, separable, ways: by deafness, and by parental language skills. In contrast to SLI, where the cause of language impairment is inherent and neurological, and therefore may or may not involve other cognitive functions, deafness directly affects children's access to spoken language but would not be expected to directly affect working memory. Nevertheless, poorer performance on working memory tasks has been noted in deaf individuals. By investigating a deaf population that includes both native and non-native signers, we can begin to explore whether concurrent impairments in working memory are linked more closely with language ability or with deafness *per se*. By avoiding tasks that require auditory instructions, stimuli and responses, we removed the immediate disadvantage that deaf children face when being compared to

hearing children. Our results indicate that when children do not have adequate exposure to a native language—regardless of its modality—this has consequences for the development of wider cognitive skills (see Campbell et al., 2014, for a discussion of the neurocognitive consequences of late age of exposure).

However, we do need to be careful when considering the data in our study—does the association between language and non-verbal working memory arise because language is mediating performance on working memory tasks concurrently, or because language has had a developmental effect on working memory up until this point in the child's life? We have been particularly careful to choose tasks that we think do not benefit from verbal mediation, and therefore, the measures should not contribute to poorer performance in this way. Nevertheless, we are aware that the nature of verbal mediation in visual tasks is not fully understood, and that research is only just beginning to explore this in children with atypical development. For example, Lidstone et al. (2012) showed that children with and without SLI were equally affected by a verbal suppression task during the Tower of London (executive memory planning task) despite the fact that children with SLI performed more poorly on the task overall. Ideally, longitudinal and training studies would help to elucidate this issue.

Furthermore, our results indicate that of the four language measures used—the Language Proficiency Profile, Expressive One Word Picture Vocabulary Test, BSL Narrative: Content and BSL Narrative: Structure—it was the Expressive One Word Picture Vocabulary Test that predicted executive-loaded working memory scores. One interpretation is that the ability to name stimuli and describe them during such tasks allows verbal mediation and draws on vocabulary skills. However, we did not have a measure of syntax, which might also be involved in verbal mediation. The grammatical structure of BSL and English is very different and not easily directly comparable, and whether syntax could be a predictor of working memory scores in our participants remains to be tested.

Finally, although we have interpreted our results as indicating that language experience directly impacts working memory, there are other differences in the developmental experience of native and non-native signers apart from language exposure that might be at play here. These include parental attachment, attention-getting strategies and social-cognitive development, among others (Marschark and Hauser, 2012). It is possible that some or all of these factors work alongside language exposure to influence working memory development. More research with larger numbers of native signers is required to fully understand these relationships.

Despite these caveats in the interpretation of our results, we argue that contrasting deaf children who grow up in optimal and suboptimal language-learning environments offers a valuable method for understanding the relationship between language and working memory. When the majority of deaf children start to develop language, they experience suboptimal conditions because the language context is predominantly oral. When this "adverse" condition is not present (i.e., when the child's deaf parent signs with them) we see a very different picture that can inform both theory and clinical practice. In particular, deafness might not be, in itself, a barrier to the development of good working memory abilities. With early exposure to an accessible sign language deaf children can demonstrate comparable skills to their hearing peers in this crucial domain. We would not wish for our results to be taken as indicating that early exposure to sign language does not help deaf children from *hearing* families—indeed, we would argue the opposite. An obvious implication for interventions with deaf children of hearing parents is for accessible language exposure to be provided early enough and in contexts where it can enhance or interact with working memory skills. A next step is to understand which aspects of language (e.g., communicative practices between interlocutors or more particular components of language such as vocabulary or syntax) are more closely involved in enabling the full development of working memory. We still lack sufficient information about the timing, amount and quality of sign language exposure that might be necessary to support age-appropriate cognitive development, and we hope to see more future research that investigates those relationships.

Acknowledgments

The study reported in this paper was funded by the Economic and Social Research Council of Great Britain (Grant 620-28-600 Deafness, Cognition and Language Research Centre). We also thank Rosalind Herman for her advice on adapting the *BSL Production Test (Narrative Skills)* into English and the *Expressive One Word Picture Vocabulary Test* into BSL.

References

Alloway, T. P., Gathercole, S. E., Adams, A.-M., and Willis, C. S. (2005). Working memory abilities in children with special educational needs. *Educ. Child Psychol.* 22, 56–67. Available online at: http://www.york.ac.uk/res/wml/Alloway%20Ed&CP.pdf

Andin, J., Orfanidou, E., Cardin, V., Holmer, E., Capek, C. M., Woll, B., et al. (2013). Similar digit-based working memory in deaf signers and hearing non-signers despite digit span differences. *Front. Psychol.* 4:942. doi: 10.3389/fpsyg.2013.00942

Avons, S., Wragg, C., Cupples, L., and Lovegrove, W. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Appl. Psycholinguist.* 19, 583–602. doi: 10.1017/S0142716400010377

Baddeley, A. D., Gathercole, S. E., and Papagno, C. (1998). The phonological loop as a language learning device. *Psychol. Rev.* 105, 158–173. doi: 10.1037/0033-295X.105.1.158

Baddeley, A. D., and Hitch, G. J. (1974). "Working memory," in *The Psychology of Learning and Motivation: Advances in Research and Theory,* Vol. 8, ed G. A. Bower (New York, NY: Academic Press), 47–89.

Bavelier, D., Newport, E. L., Hall, M., Supalla, T., and Boutla, M. (2006). Persistent differences in short-term memory span between sign and speech. *Psychol. Sci.* 17, 1090–1092. doi: 10.1111/j.1467-9280.2006.01831.x

Bavelier, D., Newport, E. L., Hall, M., Supalla, T., and Boutla, M. (2008). Ordered short-term memory differs in signers and speakers: implications for models of short-term memory. *Cognition* 107, 433–459. doi: 10.1016/j.cognition.2007.10.012

Bebko, J. M., Calderon, R., and Treder, R. (2003). The language proficiency profile-2: assessment of the global communication skills of deaf children across languages and modalities of expression. *J. Deaf Stud. Deaf Educ.* 8, 438–451. doi: 10.1093/deafed/eng034

Bebko, J. M., and McKinnon, E. E. (1993). *The Language Proficiency Profile-2*. Toronto, ON: York University.

Beer, J., Kronenberger, W. G., and Pisoni, D. B. (2011). Executive function in everyday life: implications for young cochlear implant users. *Cochlear Implants Int.* 12, S89–S91. doi: 10.1179/146701011X13001035752570

Boutla, M., Supalla, T., Newport, E. L., and Bavelier, D. (2004). Short-term memory span: insights from sign language. *Nat. Neurosci.* 7, 997–1002. doi: 10.1038/nn1298

Brownell, R. (2000). *Expressive One Word Picture Vocabulary Test, 3rd Edn*. Novato, CA: Academic Therapy Publications.

Burkholder, R. A., and Pisoni, D. B. (2003). Speech timing and working memory in profoundly deaf children after cochlear implantation. *J. Exp. Child Psychol.* 85, 63–88. doi: 10.1016/S0022-0965(03)00033-X

Campbell, R., MacSweeney, M., and Woll, B. (2014). Cochlear implantation (CI) for prelingual deafness: the relevance of studies of brain organization and the role of first language acquisition in considering outcome success. *Front. Hum. Neurosci.* 8:834 doi: 10.3389/fnhum.2014.00834

Capirci, O., Cattani, A., Rossini, P., and Volterra, V. (1998). Teaching sign Language to hearing children as a possible factor in cognitive enhancement. *J. Deaf Stud. Deaf Educ.* 3, 135–142. doi: 10.1093/oxfordjournals.deafed.a014343

Chamberlain, C., Morford, J. P., and Mayberry, R. I. (2000). *Language Acquisition by Eye*. Mahwah, NJ: Lawrence Erlbaum Associates.

Conway, C. M., Pisoni, D. B., and Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Curr. Dir. Psychol. Sci.* 18, 275–279. doi: 10.1111/j.1467-8721.2009.01651.x

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Davis, A., Bamford, J., Wilson, I., Ramkalawan, T., Forshaw, M., and Wright, S. (1997). A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment. *Health Technol. Assess.* 1, i–iv, 1–176.

Engel de Abreu, P. M. J., Abreu, N., Nikaedo, C. C., Puglisi, M. L., Tourinho, C. J., Miranda,M. C., et al. (2014). Executive functioning and reading achievement in school: a study of Brazilian children assessed by their teachers as poor readers. *Front. Psychol.* 5:550. doi: 10.3389/fpsyg.2014.00550

Fagan, M. K., Pisoni, D. B., Horn, D. L., and Dillon, C. M. (2007). Neuropsychological processes associated with vocabulary, reading, and working memory in deaf children with cochlear implants. *J. Deaf Stud. Deaf Educ.* 12, 461–471. doi: 10.1093/deafed/enm023

Faulkner, K. F., and Pisoni, D. B. (2013). Some observations about cochlear implants: challenges and future directions. *Neurosci. Discov.* 1:9. doi: 10.7243/2052-6946-1-9

Figueras, B., Edwards, L., and Langdon, D. (2008). Executive function and language in deaf children. *J. Deaf Stud. Deaf Educ.* 13, 362–377. doi: 10.1093/deafed/enm067

Gathercole, S. E., Willis, C. S., Emslie, H., and Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: a longitudinal study. *Dev. Psychol.* 28, 887–898. doi: 10.1037/0012-1649.28.5.887

Geraci, C., Gozzi, M., Papagno, C., and Cecchetto, C. (2008). How grammar can cope with limited short term memory: simultaneity and seriality in sign languages. *Cognition* 106, 780–804. doi: 10.1016/j.cognition.2007.04.014

Hall, M., and Bavelier, D. (2010). "Working memory, deafness and sign language," in *The Handbook of Deaf Studies, Language and Education*, Vol. 2, eds M. Marschark and P. E. Spencer (Oxford: Oxford University Press), 458–472.

Henry, L. A. (2001). How does the severity of a learning disability affect working memory performance? *Memory* 9, 233–247. doi: 10.1080/09658210042000085

Henry, L. A., Messer, D. J., and Nash, G. (2012). Executive functioning in children with specific language impairment. *J. Child Psychol. Psychiatry* 53, 37–45. doi: 10.1111/j.1469-7610.2011.02430.x

Herman, R., Grove, N., Holmes, S., Morgan, G., Sutherland, H., and Woll, B. (2004). *Assessing BSL Development: Production Test (Narrative Skills)*. London: City University.

Hintermair, M. (2013). EF and behavioural problems in deaf and hard-of-hearing students at general and special schools. *J. Deaf Stud. Deaf Educ.* 18, 344–359. doi: 10.1093/deafed/ent003

Hirshorn, E. A., Fernandez, N. M., and Bavelier, D. (2012). Routes to short-term memory indexing: lessons from deaf native users of American Sign Language. *Cogn. Neuropsychol.* 29, 85–103. doi: 10.1080/02643294.2012.704354

Kapa, L. L., and Colombo, J. (2014). Executive function predicts artificial language learning. *J. Mem. Lang.* 76, 237–252. doi: 10.1016/j.jml.2014.07.004

King, J., and Just, M. A. (1991). Individual differences in syntactic processing: the role of working memory. *J. Mem. Lang.* 30, 580–602. doi: 10.1016/0749-596X(91)90027-H

Kormos, J., and Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism* 11, 261–271. doi: 10.1017/S1366728908003416

Krakow, R. A., and Hanson, V. L. (1985). Deaf signers and serial recall in the visual modality: memory for signs, fingerspelling and print. *Mem. Cognit.* 13, 265–272. doi: 10.3758/BF03197689

Kwok, R. K. W., and Ellis, A. W. (2014). Visual word learning in adults with dyslexia. *Front. Hum. Neurosci.* 8:264. doi: 10.3389/fnhum.2014.00264

Kyllonen, P. C., and Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence* 14, 389–433.

Lederberg, A., and Mobley, C. E. (1990). The effect of hearing impairment on the quality of attachment and mother-toddler interaction. *Child Dev.* 61, 1596–1604. doi: 10.2307/1130767

Lev-Ari, S. (2015). Comprehending non-native speakers: theory and evidence for adjustment in manner of processing. *Front. Psychol.* 5:1546. doi: 10.3389/fpsyg.2014.01546

Lidstone, J., Meins, E., and Fernyhough, C. (2012). Verbal mediation of cognition in children with Specific Language Impairment. *Dev. Psychopathol.* 24, 651–660. doi: 10.1017/S0954579412000223

Lieberman, A. M., and Mayberry, R. I. (2015). "Studying sign language acquisition," in *Research Methods in Sign Language Studies: A Practical Guide* eds E. Orfanidou, B. Woll, and G. Morgan (Malden, MA: Wiley), 281–299.

Marschark, M., and Hauser, P. C. (2012). *How Deaf Children Learn*. Oxford: Oxford University Press.

Melby-Lervag, M., Lervag, A., Lyster, S.-A. H., Klem, M., Hagtvet, B., and Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychol. Sci.* 23, 1092–1098. doi: 10.1177/0956797612443833

Morgan, G., and Woll, B. (eds.). (2002). *Directions in Sign Language Acquisition*. Amsterdam: John Benjamins.

Parasnis, I., Samar, V., and Bettger, J. (1996). Does deafness lead to enhancement of visual spatial cognition in children? Negative evidence from deaf nonsigners. *J. Deaf Stud. Deaf Educ.* 1, 145–152. doi: 10.1093/oxfordjournals.deafed.a014288

Pickering, S., and Gathercole, S. E. (2001). *Working Memory Test Battery for Children*. London: Psychological Corporation.

Schick, B., Marschark, M., and Spencer, P. E. (eds.). (2005). *Advances in the Sign Language Development of Deaf Children*. New York, NY: Oxford University Press.

Shute, V. J. (1991). Who is likely to acquire programming skills? *J. Educ. Comput. Res.* 7, 1–24. doi: 10.2190/VQJD-T1YD-5WVB-RYPJ

Ullman, M. T., and Pierpont, E. I. (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex* 41, 399–433. doi: 10.1016/S0010-9452(08)70276-4

van der Lely, H. K. J., and Howard, D. (1993). Specifically language impaired children: linguistic impairment or short term memory deficit? *J. Speech Hear. Res.* 37, 1193–1207. doi: 10.1044/jshr.3606.1193

Wallace, G., and Corballis, M. C. (1973). Short-term memory and coding strategies in the deaf. *J. Exp. Psychol.* 99, 334–348. doi: 10.1037/h0035372

Wang, J., and Napier, J. (2013). Signed language working memory capacity of signed language interpreters and deaf signers. *J. Deaf Stud. Deaf Educ.* 18, 271–286. doi: 10.1093/deafed/ens068

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

Wechsler, D., and Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability*. San Antonio, TX: Harcourt Assessment.

Wilson, M., Bettger, J. G., Niculae, I., and Klima, E. (1997). Modality of language shapes working memory: evidence from digit span and spatial span in ASL signers. *J. Deaf Stud. Deaf Educ.* 2, 150–160. doi: 10.1093/oxfordjournals.deafed.a014321

Wilson, M., and Emmorey, K. (2006). Comparing sign language and speech reveals a universal limit on short-term memory capacity. *Psychol. Sci.* 17, 682–683. doi: 10.1111/j.1467-9280.2006.01766.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Marshall, Jones, Denmark, Mason, Atkinson, Botting and Morgan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

PERSPECTIVE
published: 27 July 2015
doi: 10.3389/fpsyg.2015.01074

Executive functions in mono- and bilingual children with language impairment – issues for speech-language pathology

Olof Sandgren and Ketty Holmström*

Department of Logopedics, Phoniatrics, and Audiology, Clinical Sciences, Lund University, Lund, Sweden

The clinical assessment of language impairment (LI) in bilingual children imposes challenges for speech-language pathology services. Assessment tools standardized for monolingual populations increase the risk of misinterpreting bilingualism as LI. This Perspective article summarizes recent studies on the assessment of bilingual LI and presents new results on including non-linguistic measures of executive functions in the diagnostic assessment. Executive functions shows clinical utility as less subjected to language use and exposure than linguistic measures. A possible bilingual advantage, and consequences for speech-language pathology practices and future research are discussed.

Keywords: bilingualism, language impairment, executive functions, bilingual advantage, speech-language pathology

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Mary Rudner,
Linköping University, Sweden
Mako Okanda,
Otemon Gakuin University, Japan
Annette Sophie Sundqvist,
Linkoping University, Sweden

****Correspondence:***
Olof Sandgren,
Department of Logopedics, Phoniatrics, and Audiology, Clinical Sciences, Lund University, S-221 85 Lund, Sweden
olof.sandgren@med.lu.se

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *19 March 2015*
Accepted: *13 July 2015*
Published: *27 July 2015*

Citation:
Sandgren O and Holmström K (2015) Executive functions in mono- and bilingual children with language impairment – issues for speech-language pathology. Front. Psychol. 6:1074. doi: 10.3389/fpsyg.2015.01074

Executive Functions in Bilingual Children

The executive functions of bilingual children have repeatedly been shown to exceed those of monolingual peers. Bilingual children outperform monolinguals on measures of inhibition, task switching, and working memory. Bialystok (1999) used a dimensional change card sort task to assess bilingual 3- to 6-year-old children's attentional control when the principle for sorting the cards changed from color to shape. The results revealed a bilingual advantage interpreted as a superior ability to inhibit incorrect responses. An ensuing experiment further traced the bilingual advantage to a specific superiority in disregarding no longer relevant information, most evident for perceptual, rather than semantic, features, and for tasks of greater complexity (Bialystok and Martin, 2004). A greater bilingual advantage in more complex tasks was also confirmed by Bialystok (2011) who showed greater performance in tasks with high demands on executive functions and on coordinating visual and auditory information. However, the bilingual advantage in attentional control extends beyond the visual domain. Using non-verbal and verbal go/no-go tasks, requiring participants to alternatingly respond to non-verbal sounds (e.g., a barking dog and a ringing bell) and verbal auditory stimuli (e.g., /pa/ and /ba/), Foy and Mann (2014) found a bilingual advantage regarding both accuracy and response times for non-verbal, but not verbal, trials.

A bilingual advantage has also been found for working memory, again with larger effects for complex tasks imposing greater executive function demands. Morales et al. (2013) hypothesized that bilingual children would exhibit better working memory as an effect of its central role in the executive functions necessary to control and coordinate two language systems. The authors contrasted congruent trials, with isolated working memory demands (remembering rules),

and incongruent trials with additional demands on executive control (remembering rules and following shifting instructions while ignoring distraction). While bilingual and monolingual 5-year-old children performed similarly on the congruent trials, with minimal demands on executive functions beyond working memory, the bilingual children responded faster. On the incongruent trials, with greater overall executive function demands, the bilingual advantage was shown by both greater accuracy and faster responses (Morales et al., 2013). Similar results had previously been presented by Carlson and Meltzoff (2008) who found a bilingual advantage for conflict tasks, similar to the incongruent trials, but not for less complex delay tasks, only requiring working memory. Furthermore, the bilingual advantage in executive functions outweighed a socio-economic disadvantage and lower language scores in the bilingual group (Carlson and Meltzoff, 2008).

To summarize, the results point to domain-general beneficial effects of bilingualism on executive functions, as further confirmed by a meta-analysis of 63 studies on the cognitive outcomes of bilingualism (Adesope et al., 2010) revealing the largest mean effect sizes for attentional control (0.96), abstract and symbolic representation (0.57), and working memory (0.48). Furthermore, the bilingual advantage grows with increasing task complexity and increasing executive function demands.

Executive Functions in Children with Language Impairment

In contrast to the advantage in executive functions evidenced by bilinguals with typical language development, monolingual children with LI have been found to be at a disadvantage compared to peers with typical language development. Im-Bolter et al. (2006) found 7- to 12-year-old children with LI to score lower than same-age peers on tasks requiring inhibition of responses and addition of information to be held in working memory. Vugs et al. (2014) found working memory deficits of 4- to 5-year-olds with LI to extend beyond the verbal domain to also include visuospatial working memory deficits, a finding taken as evidence of domain-general effects of LI with impact also on non-verbal aspects of cognition (for similar results, see Hoffman and Gillam, 2004). With 89 percent of the participants identified correctly as either LI or typically developing (TD), the authors could establish the clinical utility of working memory assessment in clinical decision making. Furthermore, using parent ratings of children's executive functions, the authors were able to document deficits in several executive functions, including inhibition (Vugs et al., 2014).

Henry et al. (2012) examined the executive functions of children with diagnosed LI in comparison to peers with undiagnosed low language/cognitive functioning, and TD. The authors found lower executive functions for participants with LI, with particular deficits in areas including verbal and non-verbal working memory, and non-verbal inhibition. Furthermore, the group difference remained significant despite adjustment for verbal IQ, indicating that the findings could not be attributed to reduced language ability. Similarly to Vugs et al. (2014), the authors found support for a domain-general impairment, and pointed to the possible clinical meaningfulness of evaluating executive functions in the assessment of LI. Furthermore, the group with undiagnosed language problems performed similarly to the group with LI on almost all measures, further supporting the clinical utility of the measures (Henry et al., 2012).

The findings of negative domain-general consequences of LI have inspired research and implementation of non-linguistic cognitive treatments to remediate the effects. While showing improvements in trained areas, establishing that executive functions are modifiable by intervention (see, e.g., Thorell et al., 2009; Holmes and Gathercole, 2014) research has yet to provide conclusive evidence of transfer to other executive functions (see, e.g., Melby-Lervåg and Hulme, 2013) or effects exceeding those of targeted language intervention (Ebert et al., 2014). However, small scale studies using single-case experimental designs have shown promising results, indicating a causal rather than merely correlational association between non-linguistic processing and language ability, in need of replication in larger samples (see, Ebert and Kohnert, 2009; Ebert et al., 2012).

Executive Functions in Bilingual Children with Language Impairment

The interaction of bilingualism and LI on executive functions remains largely unexplored. As indicated by the results above, bilingualism appears to have the potential to improve on the domain-general cognitive aspects shown to be affected by LI, and which underlie LI in theoretical constructs (see, e.g., Leonard et al., 2007, on limited processing capacity theory). If so, bilingual children with LI will present a unique linguistic and cognitive profile, distinct from those of both TD second language learners and monolinguals with LI (for a discussion, see Peets and Bialystok, 2010).

Present Study

Below, we briefly outline the aims, method, and results of an on-going study investigating a possible bilingual advantage in the executive functions of Swedish–Arabic bilingual children with LI, followed by a discussion of the implications of the results for SLP services and research.

Aims

To investigate whether bilingual Swedish–Arabic children with LI exhibit a bilingual advantage in executive functions.

Method

Fifty-four children participated in assessment of short term memory [digit span forward, WISC-IV (Wechsler, 2004); verbatim number recall], working memory [digit span backward,

WISC-IV (Wechsler, 2004), reverse order number recall], and inhibition [Berg Card Sorting Test (BCST; Mueller, 2010, sorting 128 cards according to undisclosed rules of number, color, and shape], to investigate executive functions as part of a larger study of bilingual lexical development. Prior to inclusion in the study, all participants with LI were diagnosed by a certified speech-language pathologist. Participants with TD were free from parental or teacher concern regarding language or attention. Initial analyses of receptive vocabulary, using conceptual scoring, taking into account knowledge in both languages of bilingual participants, showed equal performance between mono- and bilingual children, with and without LI, respectively (p's > 0.4). LI and TD participants were recruited from the same schools in order to reduce possible differences in socio-economic factors. Recruitment of participants and assessments were approved by the Regional Ethics Review Board for southern Sweden, approval number 2010/717.

Socio-economic status was scored from the level of parental education; primary (compulsory schooling, 1), secondary (compulsory or non-compulsory, 2), or tertiary (university level, 3) education. Arabic was the first language of both parents to all bilingual participants, and Swedish the first language of both parents to all monolingual participants. All bilingual children attended Swedish-speaking schools and had attended Swedish preschools for more than 2 years prior to the assessment. Parental reports showed the participants to be exposed to Arabic primarily at home, and to Swedish in school. No bilingual participant was reported to use either language exclusively. All participants passed a 20 dB pure-tone hearing screening at 1, 2, and 4 kHz and performed above the 10th percentile on Raven's Progressive Matrices. Mean values for participant characteristics and dependent variables are presented in **Table 1**.

Assessments of digit span forward, digit span backward and BCST were performed in accordance with the procedures described in the WISC-IV (Wechsler, 2004) and BCST (Mueller, 2010) manuals. For the bilingual participants, assessment of digit span was conducted in both Arabic and Swedish. No significant difference in performance was found [forward: $t(24) = 0.38$, $p = 0.70$; backward: $t(24) = 1.76$, $p = 0.10$] and results for Swedish are used in all subsequent analyses and discussions.

Results

The results presented here are preliminary and should be interpreted accordingly. All raw scores were converted to z-scores. Correct responses on digit span forward, digit span backward and BCST were entered as dependent variables in a multivariate ANOVA with group as the independent variable. A statistically significant difference between the groups was found for an overall measure of executive functions, combining the scores of all dependent variables [$F(9,150) = 4.12$, $p < 0.001$, Pillai's Trace $= 0.60$, $\eta_p^2 = 0.20$]. The group difference remained significant when the dependent variables were analyzed separately [digit span forward; $F(3,50) = 11.46$, $p < 0.001$, $\eta_p^2 = 0.41$, digit span backward; $F(3,50) = 7.31$, $p < 0.001$, $\eta_p^2 = 0.31$, BCST; $F(3,50) = 4.93$, $p = 0.004$, $\eta_p^2 = 0.23$; see **Figure 1** and **Table 1**]. *Post hoc* analyses with LSD revealed BLI to perform on par with MLI on all measures [digit span forward; $p = 0.12$, $d = 0.96$, digit span backward; $p = 0.27$, $d = 0.36$, BCST; $p = 0.45$, $d = 0.28$]. MTD outperformed BTD on digit span forward ($p = 0.01$, $d = 0.81$) while similar performance between TD groups was found for digit span backward ($p = 0.60$, $d = 0.24$) and BCST ($p = 0.97$, $d = 0.01$). For comparisons between LI and TD groups, BLI and BTD performed similarly on digit span forward ($p = 0.13$, $d = 0.61$) while BLI performed significantly below BTD on digit span backward ($p = 0.02$, $d = 1.10$) and BCST ($p = 0.03$, $d = 0.78$). MTD outperformed MLI on all measures [digit span forward; $p < 0.001$, $d = 2.65$, digit span backward; $p < 0.001$, $d = 1.23$, BCST; $p = 0.003$, $d = 1.29$].

To summarize, BLI and MLI performed on par on all dependent variables, while BTD and MTD differed only on digit span forward. BLI differed from BTD peers on digit span backward and BCST while MLI differed significantly from MTD on all measures. Digit span backward and digit span forward produced the largest effect sizes for BLI-BTD and MLI-MTD comparisons, respectively. For BLI-MLI and BTD-MTD comparisons, BCST produced the smallest effect sizes.

Discussion

While preliminary, the results replicate earlier findings which indicate that measures of non-linguistic processing may provide

TABLE 1 | Mean values for participant characteristics and dependent variables.

	BLI ($n = 9$) *M* (SD)	BTD ($n = 18$) *M* (SD)	MLI ($n = 9$) *M* (SD)	MTD ($n = 18$) *M* (SD)	*p*
Age	6;10 (0;7)	7;0 (0;7)	6;10 (0;8)	6;11 (0;6)	0.75
SES	1.79 (0.49)	2.25 (0.59)	2.37 (0.52)	2.44 (0.63)	0.10
Arabic exposure (%)	53.0 (5.2)	50.9 (9.3)			0.59
Arabic use (%)	37.3 (15.3)	42.3 (14.6)			0.51
Digit span forward	−0.46 (0.58)	0.03 (1.0)	−1.06 (0.65)	0.73 (0.69)	<0.001
Digit span backward	−0.52 (0.89)	0.30 (0.56)	−0.97 (1.50)	0.45 (0.63)	<0.001
Berg Card Sorting Test (BCST)	−0.49 (1.17)	0.33 (0.92)	−0.82 (1.12)	0.32 (0.57)	0.004

BLI, Bilingual LI; BTD, Bilingual TD; MLI, monolingual LI; MTD, monolingual TD.

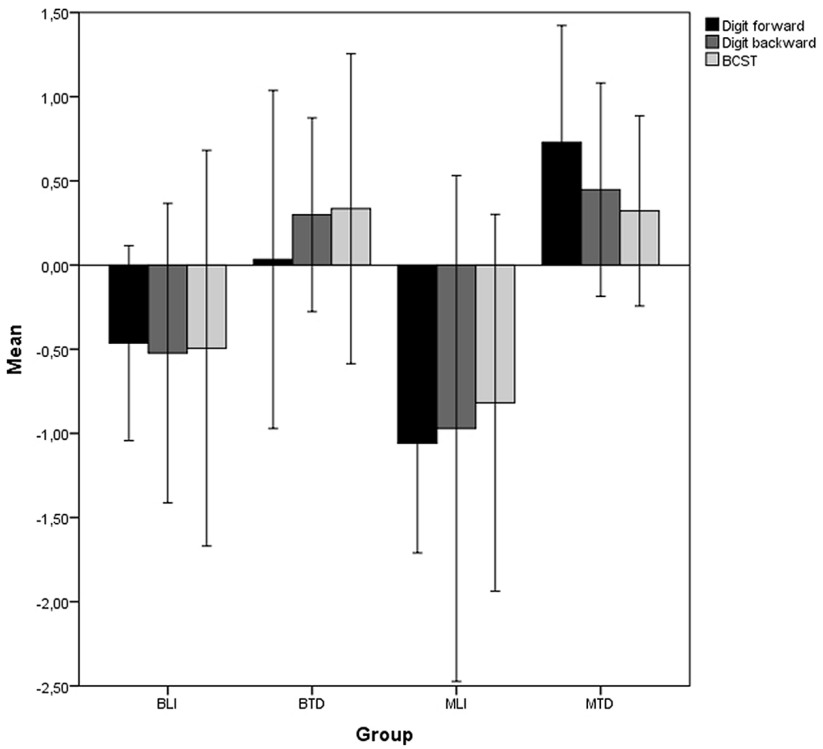


FIGURE 1 | *Z*-scores with ±1 SD error bars for dependent variables (digit span forward, digit span backward, Berg Card Sorting Test). BLI, Bilingual LI; BTD, Bilingual TD; MLI, monolingual LI; MTD, monolingual TD.

important information in multilingual contexts (Paradis, 2010a,b). The study fails to provide evidence for a bilingual advantage in bilingual children with LI. Importantly, a bilingual disadvantage is also absent, somewhat surprisingly considering lower socio-economic status and lower Swedish language exposure for the bilingual than for the monolingual groups. The effects and interactions of socio-economic status (previously shown to attenuate a bilingual advantage in executive functions, see Morton and Harper, 2007), language proficiency (shown to affect cognitive processing in younger children, see Okanda et al., 2010), task complexity in relation to LI, and sample size may all play a role in explaining the absent bilingual advantage. While linguistic measures are commonly found to differ between mono- and bilingual children, equal performance in the present study indicates that executive functions are less subjected to influence from language exposure. Still, the measures appear to tap linguistic processing. For digit span forward, measuring short term memory, the best performance is found in monolinguals with TD, and the measure is also the best to separate monolinguals with and without LI. Interestingly, the bilinguals with and without LI show equal performance in digit span forward, a finding which could, as suggested by Morales et al. (2013), be interpreted as a bilingual advantage. The task of repeating digits may be complex enough to evoke an advantage for the bilinguals with LI, while their TD peers, with overall greater linguistic abilities, will not find the task challenging enough. In contrast, digit span backward, measuring working memory, appears to evoke an advantage also for bilinguals with TD, more clearly separating the bilingual children with and without LI for this measure.

The results of these preliminary analyses indicate that the clinical benefits of including executive functions in the assessment of LI are limited, at least in terms of identifying children with LI. Our sample is small, and replication is needed to see which results can be generalized. Subsequent studies should further investigate the influence of language proficiency on a bilingual advantage in executive functions. As suggested by Peets and Bialystok (2010), second language learners early in development may not show the effect, or show a bilingual advantage in other tasks than peers with more developed linguistic capacities. If LI is the result of atypical cognitive processes affecting, for example, executive functions, bilingualism might offset these processes, and improve language development. However, all children with LI may not exhibit deficits in executive functions, and further analyses must delve deeper into the interaction between executive functions and language ability, by investigating the individual language profile of participants with differences in executive functions. This may enable more individualized intervention, as well as improved differential diagnostics in speech-language pathology. For example, this may help determine the threshold in executive functions necessary for positive effects on language outcome, and contribute to a better understanding of the complex cognitive and language profiles of bilingual children with LI.

References

Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644. doi: 10.1111/1467-8624.00046

Bialystok, E. (2011). Coordination of executive functions in monolingual and bilingual children. *J. Exp. Child Psychol.* 110, 461–468. doi: 10.1016/j.jecp.2011.05.005

Bialystok, E., and Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Sci.* 7, 325–339. doi: 10.1111/j.1467-7687.2004.00351.x

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x

Ebert, K. D., and Kohnert, K. (2009). Non-linguistic cognitive treatment for primary language impairment. *Clin. Linguist. Phon.* 23, 647–664. doi: 10.1080/02699200902998770

Ebert, K. D., Kohnert, K., Pham, G., Disher, J. R., and Payesteh, B. (2014). Three treatments for bilingual children with primary language impairment: examining cross-linguistic and cross-domain effects. *J. Speech Lang. Hear. Res.* 57, 172–186. doi: 10.1044/1092-4388(2013/12-0388)

Ebert, K. D., Rentmeester-Disher, J., and Kohnert, K. (2012). Nonlinguistic cognitive treatment for bilingual children with primary language impairment. *Clin. Linguist. Phon.* 26, 485–501. doi: 10.3109/02699206.2012.660226

Foy, J. G., and Mann, V. A. (2014). Bilingual children show advantages in nonverbal auditory executive function task. *Int. J. Biling.* 18, 717–729. doi: 10.1177/1367006912472263

Henry, L. A., Messer, D. J., and Nash, G. (2012). Executive functioning in children with specific language impairment. *J. Child Psychol. Psychiatry* 53, 37–45. doi: 10.1111/j.1469-7610.2011.02430.x

Hoffman, L. M., and Gillam, R. B. (2004). Verbal and spatial information processing constraints in children with specific language impairment. *J. Speech Lang. Hear. Res.* 47, 114–125. doi: 10.1044/1092-4388(2004/011)

Holmes, J., and Gathercole, S. E. (2014). Taking working memory training from the laboratory into schools. *Educ. Psychol.* 34, 440–450. doi: 10.1080/01443410.2013.797338

Im-Bolter, N., Johnson, J., and Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Dev.* 77, 1822–1841. doi: 10.1111/j.1467-8624.2006.00976.x

Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., and Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *J. Speech Lang. Hear. Res.* 50, 408–428. doi: 10.1044/1092-4388(2007/029)

Melby-Lervåg, M., and Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Dev. Psychol.* 49, 270–291. doi: 10.1037/a0028228

Morales, J., Calvo, A., and Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *J. Exp. Child Psychol.* 114, 187–202. doi: 10.1016/j.jecp.2012.09.002

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x

Mueller, S. T. (2010). *PEBL: The Psychology Experiment Building Language (Version 0.11)*. Available at: http://pebl.sourceforge.net

Okanda, M., Moriguchi, Y., and Itakura, S. (2010). Language and cognitive shifting: evidence from young monolingual and bilingual children. *Psychol. Rep.* 107, 68–78. doi: 10.2466/03.10.28.PR0.107.4.68-78

Paradis, J. (2010a). The interface between bilingual development and specific language impairment. *Appl. Psycholinguist.* 31, 227–252. doi: 10.1017/S0142716409990373

Paradis, J. (2010b). Response to Commentaries on the interface between bilingual development and specific language impairment. *Appl. Psycholinguist.* 31, 345–362. doi: 10.1017/S0142716409990531

Peets, K. F., and Bialystok, E. (2010). An integrated approach to the study of specific language impairment and bilingualism. *Appl. Psycholinguist.* 31, 315–319. doi: 10.1017/S0142716409990488

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., and Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Dev. Sci.* 12, 106–113. doi: 10.1111/j.1467-7687.2008.00745.x

Vugs, B., Hendriks, M., Cuperus, J., and Verhoeven, L. (2014). Working memory performance and executive function behaviors in young children with SLI. *Res. Dev. Disabil.* 35, 62–74. doi: 10.1016/j.ridd.2013.10.022

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children*, 4th Edn. London: Pearson Assessment.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Sandgren and Holmström. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

MINI REVIEW
published: 24 June 2015
doi: 10.3389/fpsyg.2015.00871

On the interaction of speakers' voice quality, ambient noise and task complexity with children's listening comprehension and cognition

Viveka Lyberg-Åhlander, K. J. Brännström and Birgitta S. Sahlén*

Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University, Lund, Sweden

OPEN ACCESS

Edited by:
Mary Rudner, Linköping University, Sweden

Reviewed by:
Karen A. Gordon, The Hospital for Sick Children, Canada
Suzanne Carolyn Purdy, University of Auckland, New Zealand

****Correspondence:***
Viveka Lyberg-Åhlander, Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University Hospital, S-221 85 Lund, Sweden
viveka.lyberg_ahlander@med.lu.se

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *23 February 2015*
Accepted: *12 June 2015*
Published: *24 June 2015*

Citation:
Lyberg-Åhlander V, Brännström KJ and Sahlén BS (2015) On the interaction of speakers' voice quality, ambient noise and task complexity with children's listening comprehension and cognition. Front. Psychol. 6:871. doi: 10.3389/fpsyg.2015.00871

Suboptimal listening conditions interfere with listeners' on-line comprehension. A degraded source signal, noise that interferes with sound transmission, and/or listeners' cognitive or linguistic limitations are examples of adverse listening conditions. Few studies have explored the interaction of these factors in pediatric populations. Yet, they represent an increasing challenge in educational settings. We will in the following report on our research and address the effect of adverse listening conditions pertaining to speakers' voices, background noise, and children's cognitive capacity on listening comprehension. Results from our studies clearly indicate that children risk underachieving both in formal assessments and in noisy class-rooms when an examiner or teacher speaks with a hoarse (dysphonic) voice. This seems particularly true when task complexity is low or when a child is approaching her/his limits of mastering a comprehension task.

Keywords: comprehension, voice, noise, cognition, children

Background

Poor listening environments are challenging for typically developing children with normal hearing and even more so for children struggling with listening comprehension in different disability groups (Khalfa et al., 2004). Noise that interferes with sound transmission, forces students to allocate cognitive capacity to suppress the task irrelevant input. This allocation spares less capacity for the processing and recall of the content (Shield and Dockrell, 2008; Sörqvist, 2010; Klatte et al., 2013). However, little attention has been paid to the role which source signal alterations, for example changes in speakers' speaking rate, may play for children's listening comprehension. In one of our studies, 8-year-olds listened to recorded sentences read aloud by a speech language pathologist speaking with either fast, normal or slow speech rate (Haake et al., 2014). The slower speech rate was generally associated with better performance on a language comprehension test. Children with stronger working memory capacity (WMC) benefitted more from slow speech rate than their peers, but only for more complex sentences. The slower speech rate did not improve performance on the more complex tasks in children with weaker WMC,

probably because these tasks were beyond their grasp. It was concluded by the authors that it is when the child is just about to master a comprehension task that slower speech is beneficial.

The Influence of Adverse Voice Quality on Listening Comprehension

Alterations of speech rate may degrade the source signal but the risk for degradation is higher when a speaker speaks with dysphonic voice or a non-native accent (Mattys et al., 2012). A dysphonic (coml. hoarse) voice is defined as a voice that qualitatively may deviate from the 'typical' in a number of ways, e.g., pressed (hyperfunctional), breathy, rough and/or instable. The cause is an organically or functionally impaired voice function. Only a couple of studies have investigated the impact of voice quality on listening comprehension (Morton and Watson, 2001; Rogerson and Dodd, 2005). In spite of small differences in methodology, the authors' conclusions are convergent: a dysphonic teacher-voice hampers children's comprehension and listeners may judge dysphonic voices more negatively than typical voices with possible effects on motivation and learning (Morton and Watson, 2001).

Our own studies corroborate these findings and extend existing knowledge in some explorative and experimental studies. More specifically, we studied the impact of teachers' voice quality on children's accuracy, reaction times in a listening comprehension task with increasing complexity. We further studied the children's subjective experience of the voice. The experiments were performed either in silence or in background babble-noise (Brännström et al., 2014; Lyberg-Åhlander et al., 2015a,b).

We used a digitalized version of a language comprehension test, the TROG-2 (Bishop, 2003, 2009), which is a picture selection test consisting of 80 sentences, organized into 20 blocks with increasing lingusitic complexity. Accuracy, self-corrections and speed (response times) were measured. To assess WMC, the Competing Language Processing Task (CLPT; Gaulin and Campbell, 1994), was used. The CLPT is a test used for assessment of complex WMC. In the CLPT, initially the participant is asked to judge the semantic acceptability of a sentence and thereafter, in blocks of 1–6 sentences (a total of 42 sentences), they are asked to repeat the final words of each sentence. To assess executive functioning the Elithorn's Mazes (EM, WISC–IV; Wechsler, 2004) were used. In all four studies reported below, we utilized a between-group design. The children listened to the recorded sentences read by the same female speaker, either using her normal voice or a dysphonic voice, either mimicked or induced through vocal loading. In each study, around 90 typically developing normal hearing 8-year olds from schools in Southern Sweden were included.

The first study by Lyberg-Åhlander et al. (2015a) was performed with a mimicked dysphonic voice and no ambient noise. We found no overall effect of the mimicked and moderately dysphonic voice on comprehension. However, the children listening to the dysphonic voice achieved significantly lower TROG-2 scores for sentences in the more complex blocks of the test ("the man but not the horse is jumping"). These children also made significantly more self-corrections than those listening to the typical voice, but this was restricted to the less complex sentences ("the girl is sitting"). Decreased accuracy in more complex tasks was interpreted as indicating that the mimicked dysphonic speaker's voice forced children to allocate capacity to the processing of the voice signal at the expense of listening comprehension, particularly when the linguistic difficulty is of borderline complexity for the child. The scores on EM correlated significantly to the TROG-2 results. We also analyzed response times. Response time is often used as measures for listening effort in adults and are, by some researchers, considered a reliable measure for listening effort in children (Hick and Tharpe, 2002). Preliminary analyses yielded no overall difference between voice qualities, but response times increased with task difficulty in both conditions and were longer for girls in the dysphonic condition (with mimicked and vocally loading induced dysphonia) as compared to the girls in the typical voice condition and to the boys in both conditions. Based on our data we believe that several other factors such as interest, motivation, and socio-cultural aspects underpin response times.

The Combined Effect of a Dysphonic Voice Quality and Noise on Comprehension

In yet another study, Lyberg-Åhlander et al. (2015b) explored what happens when children listen to a typical versus a dysphonic speaker in simultaneous background babble-noise. Speaking in a noisy environment will also change the voice quality of a speaker with a typical voice. Therefore, the voice-paradigm had to be altered to achieve two ecologically valid voice qualities. The female speaker was now recorded as she was making herself heard while speaking in babble-noise. During the study, one group of children listened to the speaker recorded with her somewhat strained but 'typical' voice in babble-noise (Holube et al., 2010) and another group listened to her dysphonic voice, which was induced by a vocal loading task before the recording. The vocal loading task refers to when the speaker was asked to read out loud for 30 min in 85dB babble-noise (Whitling et al., 2015). This mode of vocal loading, common in noisy classrooms, often causes a speaker with a healthy voice to raise the fundamental frequency and to use a more hyperfunctional phonation. Speaking over noise changes the spectrum of the voice as compared to the typical voice, and may result in an increase or decrease of noise in the higher part of the spectrum. The ecological validity of the voice qualities (typical/dysphonic) was assessed by an expert panel where the dysphonic voice was judged as significantly more disordered.

The TROG-2 results did not differ between the groups. We concluded that the background babble-noise, present in both conditions, might have masked a possible additional effect of the dysphonic voice. However, significant differences between voice conditions were found for the interaction between WMC and linguistic task-complexity, particularly in tasks representing

intermediate difficulty. In the dysphonic voice condition, children with stronger WMC scored significantly higher on easier blocks, whereas, in the typical voice condition the cognitively stronger children scored higher on more difficult blocks.

Unfortunately, a direct comparison between the results of these two studies Lyberg-Åhlander et al. (2015a,b), is impeded by differences in transducers used to present the voices and by the use of mimicked versus authentic dysphonia. Therefore, the relative contribution of the voice quality *per se* cannot be teased out. Even so, importantly, these combined results indicate synergistic detrimental effects on children's listening comprehension in a class-room when dysphonic teachers try to make themselves heard in ambient noise.

The Interaction of Perceptual Load, Task Complexity and Attitude to Voice

Some of the results from these studies are complex and at first counterintuitive. For instance, why should a dysphonic voice lead children to make more self-corrections on easier tasks than on more difficult tasks? According to the perceptual load theory (Lavie, 2005), sufficiently easy tasks free cognitive capacity to process task-irrelevant stimuli in adults. This may explain the increased amount of self-corrections in the easier tasks in the dysphonic condition in the earlier study (Lyberg-Åhlander et al., 2015a). The children may have had the cognitive capacity needed to process, or even to get disturbed by, the dysphonic voice. Results in the later study by Lyberg-Åhlander et al. (2015b) may be explained accordingly. In this study, children with stronger WMC, performed better on the more difficult tasks when listening to the typical voice in noise (i.e., lower perceptual load and higher cognitive complexity) and on the easier tasks when listening to the dysphonic voice in noise (i.e., higher perceptual load and lower cognitive complexity). Detrimental effects of adverse conditions on listening comprehension may thus decrease when perceptual load increases, as was the case for children with stronger WMC. This is in line with the perceptual load hypothesis stating that, in adults, the effect of task-irrelevant stimuli diminishes when the task itself is sufficiently complex.

It has previously been suggested that negative attitudes toward a teacher's voice may influence the teacher–child relation and as a consequence may influence motivation and learning outcomes negatively (Morton and Watson, 2001). In Brännström et al. (2014), we therefore investigated children's subjective ratings of the speakers' voices using data from Lyberg-Åhlander et al. (2015b). Children thus listened to the same speaker using typical voice or with vocal-loading induced dysphonic voice in ambient babble-noise. Self-reports from the children of perceived effort and attitude to the teacher voice were collected after the listening comprehension task. The children's judgments were collected with the help of emoticons, later transformed to a five-step Lickert scale. The dysphonic voice, as expected, received lower ratings compared to the ratings of the typical voice. Example children's opinions were that the speaker with the dysphonic voice was 'stressed' or 'nice but determined.' Children in the typical voice group who made more positive ratings of the voice, performed better on earlier items in the TROG-2. Accordingly, the perception of the voice related to the child's performance for low complexity tasks. Self-assessments in a pediatric population are problematic for a range of reasons and further studies are needed. Children may rate both their own and other's behavior in relation to their self-efficacy, to their own task performance and to other contextual circumstances, especially when made in hind-sight. Children might also try to either deceive or please the test-leader (DeRight and Carone, 2015).

A Developmental Perspective on Human Voice Recognition

During adverse listening conditions, whether the origin is related to the speaker, the environment or the listener, compensatory mechanisms emerge, and recalibration takes place in 'the human speech recognizer.' Memory representations of talkers' voices are stored in long-term memory (Mattys et al., 2012). A developmental perspective of this type of perceptual learning in talker recognition has been proposed by Creel and Jimenez (2012). According to these authors, young preschool children, with typical cognitive and linguistic development will cease to filter out acoustical cues during development and successively internalize such cues and finally become efficient at talker recognition and understanding as adults. This developmental perspective suggests that differences in adaptation to speakers' voice quality could be related to the child's cognitive capacity. With our between-group we can only speculate that the children with a stronger cognitive capacity and better listening comprehension may have developed more stable talker-templates. They would perhaps, as a result, be less disturbed than cognitively less mature children by a mismatch between a degraded talker signal (such as when their teacher suddenly becomes dysphonic), and their memory representations of the speaker's 'normal' voice.

Implications for Future Studies

We have recently taken several steps to reach higher ecological validity in on-going studies. As for the interaction of noise and voice, in Lyberg-Åhlander et al. (2015b), we aimed to simulate an actual classroom situation by using multi-talker babble International Speech Test Signal (ISTS; Holube et al., 2010), with six female voices constituting the noise source. Our choice of speaker and babble-noise was inspired by Zekveld et al. (2014), who conclude that speech recognition is more influenced when the disturbing signal is produced by a person of the same gender as the speaker of the target signal and, that the cognitive load is greater. This is especially true when the disturbing signal is derived from a source that is spatially close to the target signal. Our choice of a non-semantic babble-noise may, however, have made the comprehension task somewhat easier compared to if the babble would have been possible to understand (Rosen et al., 2013). Further studies will therefore utilize semantic babble.

In current studies we are addressing effects of suboptimal listening conditions on long-term memory. It is possible that the influence of voice quality on performance and attitude will change if children are assessed after a period of time when long-term memory integration has occurred. Thus a measurement that is not restricted to comprehension of sentences but that includes also comprehension of narratives both in direct connection to the task and after a period of time, could investigate the effects of episodic memory. Further, multimodal aspects of comprehension and memory in adverse conditions are explored by the use of a virtual teacher agent. This enables the systematic study of visual versus audio-visual aspects of comprehension. Using a mixture of techniques (optical markers and infrared 3D-gitter, Dutta, 2012; Gonzalez-Jorge et al., 2013) we can record both macro- (postures, gestures) and micro-level (eye blinks and lips) movements and map them onto a digital 3D-character. A virtual teacher allows further experimental control of visual aspects (sex/gender, age, clothing, etc.) as well as postural movement and gestures (amplitude, velocity, synchronization, etc.) in combination with controlled voice recordings.

Conclusion

Today, assessment and intervention in children with language, hearing, and/or cognitive impairments are increasingly based on knowledge of how cognitive functioning and acoustic processing interact. There is, however, an apparent lack of knowledge on how noise interacts with these factors. Environmental noise not only influences children's comprehension but also teacher's voices. Voice problems reach a point-prevalence of thirteen percent in Swedish teachers (Lyberg-Åhlander et al., 2011) and a career prevalence close to 60% (Roy et al., 2004). The summary of our results indicates that children risk underachievement in both formal assessments and in noisy class-rooms if an examiner or teacher speaks with a dysphonic voice, particularly when tasks demands are too low or when the child is approaching her/his limits of mastering a comprehension task. Our studies indicate that individual variations in cognitive capacity must be taken into consideration in research on the interaction of task complexity and on adverse listening conditions pertaining to the speaker and the noise environment.

Acknowledgments

We are grateful for financial support and valuable input from collaborators in the Linnaeus environments Cognition Communication and Learning (CCL), Lund University. We also thank members in the Linnaeus environment Hearing and Deafness (HEAD), Linköping University, Sweden, for fruitful discussions.

References

Bishop, D. (2003). *Test for Reception of Grammar, Version 2* (*TROG 2*). London: Pearson Assessment.

Bishop, D. (2009). *Test for Reception of Grammar, Version 2* (*TROG 2*). *Swedish version*. London: Pearson Assessment.

Brännström, K. J., Holm, L., Lyberg-Åhlander, V., Haake, M., Kastberg, T., and Sahlén, B. (2014). Children's subjective ratings and opinions of typical and dysphonic voice after performing a language comprehension task in background noise. *J. Voice.* doi: 10.1016/j.jvoice.2014.11.003 [Epub ahead of print].

Creel, S. C., and Jimenez, S. R. (2012). Differences in talker recognition by preschoolers and adults. *J. Exp. Child. Psychol.* 113, 487–509. doi: 10.1016/j.jecp.2012.07.007

DeRight, J., and Carone, D. A. (2015). Assessment of effort in children: a systematic review. *Child Neuropsychol.* 21, 1–24. doi: 10.1080/09297049.2013.864383

Dutta, T. (2012). Evaluation of the Kinect^{TM} sensor for 3-D kinematic measurement in the workplace. *Appl. Ergon.* 43, 645–649. doi: 10.1016/j.apergo.2011.09.011

Gaulin, C. A., and Campbell, T. F. (1994). Procedure for assessing verbal working memory in normal school-age children: some preliminary data. *Percept. Mot. Skills* 79(Pt 1), 55–64. doi: 10.2466/pms.1994.79.1.55

Gonzalez-Jorge, H., Riveiro, B., Vazquez-Fernandez, E., Martínez-Sánchez, J., and Arias, P. (2013). Metrological evaluation of microsoft kinect and asus xtion sensors. *Measurement* 46, 1800–1806. doi: 10.1016/j.measurement.2013.01.011

Haake, M., Hansson, K., Gulz, A., Schotz, S., and Sahlen, B. (2014). The slower the better? Does the speaker's speech rate influence children's performance on a language comprehension test? *Int. J. Speech Lang. Pathol.* 16, 181–190. doi: 10.3109/17549507.2013.845690

Hick, C. B., and Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *J. Speech Lang. Hear Res.* 45, 573–584. doi: 10.1044/1092-4388(2002/046)

Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (2010). Development and analysis of an international speech test signal (ISTS). *Int. J. Audiol.* 49, 891–903. doi: 10.3109/14992027.2010.506889

Khalfa, S., Bruneau, N., Roge, B., Georgieff, N., Veuillet, E., Adrien, J. L., et al. (2004). Increased perception of loudness in autism. *Hear. Res.* 198, 87–92. doi: 10.1016/j.heares.2004.07.006

Klatte, M., Bergstroem, K., and Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Front. Psychol.* 4:578. doi: 10.3389/fpsyg.2013.00578

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends Cogn. Sci.* 9, 75–82. doi: 10.1016/j.tics.2004.12.004

Lyberg-Åhlander, V., Haake, M., Brännström, J., Schötz, S., and Sahlén, B. (2015a). Does the speaker's voice quality influence children's performance on a language comprehension test? *Int. J. Speech Lang. Pathol.* 1, 63–73. doi: 10.3109/17549507.2014.898098

Lyberg-Åhlander, V., Holm, L., Kastberg, T., Haake, M., Brännström, K. J., and Sahlen, B. (2015b). Are children with stronger cognitive capacity more or less disturbed by classroom noise and dysphonic teachers? *Int. J. Speech Lang. Pathol.* 1–12. [Epub ahead of print].

Lyberg-Åhlander, V., Rydell, R., and Löfqvist, A. (2011). Speaker's comfort in teaching environments: voice problems in Swedish teaching staff. *J. Voice* 25, 430–440. doi: 10.1016/j.jvoice.2009.12.006

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Proc.* 27, 953–978. doi: 10.1080/01690965.2012.705006

Morton, V., and Watson, D. (2001). The impact of impaired vocal quality on children's ability to process spoken language. *Logoped. Phoniatr. Vocol.* 26, 17–25. doi: 10.1080/140154301300109080

Rogerson, J., and Dodd, B. (2005). Is there an effect of dysphonic teachers' voices on children's processing of spoken language? *J. Voice* 19, 47–60. doi: 10.1016/j.jvoice.2004.02.007

Rosen, S., Souza, P., Ekelund, C., and Majeed, A. A. (2013). Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *J. Acoust. Soc. Am.* 133, 2431–2443. doi: 10.1121/1.4794379

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., and Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *J. Speech Lang. Hear Res.* 47, 281–293. doi: 10.1044/1092-4388 (2004/023)

Shield, B. M., and Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *J. Acoust. Soc. Am.* 123, 133–144. doi: 10.1121/1.2812596

Sörqvist, P. (2010). The role of working memory capacity in auditory distraction: a review. *Noise Health* 12, 217–224. doi: 10.4103/1463-1741.70500

Wechsler, D. (2004). *WISC IV Integrated. Wechsler Intelligence Scale for Children London*, UK: Pearson Assessment.

Whitling, S., Rydell, R., and Lyberg-Åhlander, V. (2015). Design of a clinical vocal loading test with long-time measurement of noice. *J. Voice* 261, e13–e27. doi: 10.1016/j.jvoice.2014.07.012

Zekveld, A. A., Rudner, M., Kramer, S. E., Lyzenga, J., and Rönnberg, J. (2014). Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Front. Neurosci.* 8:88. doi: 10.3389/fnins.2014.00088

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Lyberg-Åhlander, Brännström and Sahlén. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PSYCHOLOGY

MINI REVIEW ARTICLE
published: 09 March 2015
doi: 10.3389/fpsyg.2015.00242

Working memory and referential communication—multimodal aspects of interaction between children with sensorineural hearing impairment and normal hearing peers

Olof Sandgren*, Kristina Hansson and Birgitta Sahlén

Department of Logopedics, Phoniatrics, and Audiology, Clinical Sciences, Lund University, Lund, Sweden

Edited by:
Mary Rudner, Linköping University, Sweden

Reviewed by:
Chloe Marshall, University College London Institute of Education, UK
Nicole Muller, Linköping University, Sweden

****Correspondence:***
Olof Sandgren, Department of Logopedics, Phoniatrics, and Audiology, Clinical Sciences, Lund University, Lasarettsgatan 21, S-221 85 Lund, Sweden
e-mail: olof.sandgren@med.lu.se

Whereas the language development of children with sensorineural hearing impairment (SNHI) has repeatedly been shown to differ from that of peers with normal hearing (NH), few studies have used an experimental approach to investigate the consequences on everyday communicative interaction. This mini review gives an overview of a range of studies on children with SNHI and NH exploring intra- and inter-individual cognitive and linguistic systems during communication. Over the last decade, our research group has studied the conversational strategies of Swedish speaking children and adolescents with SNHI and NH using referential communication, an experimental analog to problem-solving in the classroom. We have established verbal and non-verbal control and validation mechanisms, related to working memory capacity and phonological short term memory. We present main findings and future directions relevant for the field of cognitive hearing science and for the clinical and school-based management of children and adolescents with SNHI.

Keywords: child hearing impairment, referential communication, working memory, phonological short term memory, gaze behavior

LANGUAGE AND COMMUNICATION IN CHILDREN WITH SNHI

LANGUAGE

The language development of children with sensorineural hearing impairment (SNHI) with hearing aids and/or cochlear implants has, at a group level, repeatedly been shown to depart from the typical trajectory. Several studies have found approximately half of preschool children with SNHI to exhibit substantial language problems, as compared to approximately 5% in the general population (Gilbertson and Kamhi, 1995; Hansson et al., 2004; Sahlén and Hansson, 2006), with particular deficits in phonological processing (Briscoe et al., 2001; Sahlén et al., 2004; Wake et al., 2006; Wass et al., 2008) and vocabulary (Mayne et al., 1998a,b; Hansson et al., 2004), whereas results are mixed regarding grammar (Norbury et al., 2001; Hansson et al., 2007). While basic language skills can normalize with age, children with SNHI have been found not to close the gap to normal hearing (NH) peers regarding complex language functioning, for example, oral and written narrative ability (Asker-Árnason et al., 2010; Reuterskiöld et al., 2010). Intrinsic (cognitive) and extrinsic (audiological and linguistic intervention, quality and quantity of input, feedback and teaching) factors, in complex interaction, likely contribute to the substantial heterogeneity in language outcome.

COMMUNICATION

Whereas the primary purpose of language is communication, language ability—at least narrowly defined as the capacity to form linguistically coherent messages—is merely one tool necessary for successful communication. Verbal and non-verbal modalities are integrated with contextual factors to shape our ability to interact with others (Perkins, 2007). Interlocutors continuously merge the verbal message with information gathered from the partner's speech, voice, posture, field of vision, gaze direction and gestures, as well as contextual information, for example, knowledge of the world, the context and the topic of the conversation. Consequently, intra- and inter-individual linguistic, cognitive, and socio-cognitive systems interact in communication. A hearing impairment may lead to misallocation of resources with negative effects on listening ability and understanding.

While studies often include protocols or checklists considered to capture social and communicative abilities there are surprisingly few experimental studies of children with SNHI interacting with others in everyday communicative settings. Most et al. (2010) analyzed aspects of pragmatic ability in 6 to 9 year-old children with severe-to-profound SNHI (using hearing aids and/or cochlear implants) from video recorded spontaneous conversation with a speech-language pathologist. Although not consistently impaired, the children with SNHI showed particular

problems continuing the topic of the partner, and adding relevant information. Most et al. (2010) argued that the problems observed in the children with SNHI are caused by a delayed language development and limited linguistic input, resulting in an inexperience with various pragmatic behaviors and restricted perspective-taking. Compatible with a delayed language development, Toe and Paatsch (2010) presented results showing 7 to 12 year-old children with mild-to-profound SNHI to request repetition and clarification of questions to a significantly higher extent than NH peers. Similar results have been presented by our own research group using referential communication tasks, first introduced by Glucksberg and Krauss (1967), providing a compromise between experimental control and ecological validity, and designed to tap the communicative ability used in everyday activities such as giving instructions, describing things or events to a listener, and asking questions. In our studies, the referential communication tasks were designed to resemble communication between peers in structured classroom activities, rather than spontaneously occurring interaction.

REFERENTIAL COMMUNICATION—METHODOLOGY

Apart from providing details on typical communicative development, studies of referential communication have added to our knowledge of the communicative competence of individuals with a range of disabilities. In a referential communication task, the speaker is provided with an array of referents (pictures or physical objects), arranged in a predetermined pattern. The speaker's task is to describe each picture/object, and its position, to enable the listener to arrange his/her array in the same way. Referential communication tasks allow investigation of the participants' ability to produce (when in the "speaker" role), perceive and understand (when in the "listener" role) spoken messages (see **Figure 1**). Specifically, the task seeks to investigate whether the speaker can form contextually relevant messages, providing the listener with necessary information, without providing unnecessary details. The listener is evaluated on the ability to detect and resolve ambiguities through his/her use of questions. If, for example, the speaker describes a picture of a face as "It's a man with a beard" this would provide sufficient information if all other referents lacked these characteristics. However, if the competing referents included other men with beards the listener would have to request additional information, for example "Is he wearing glasses?"

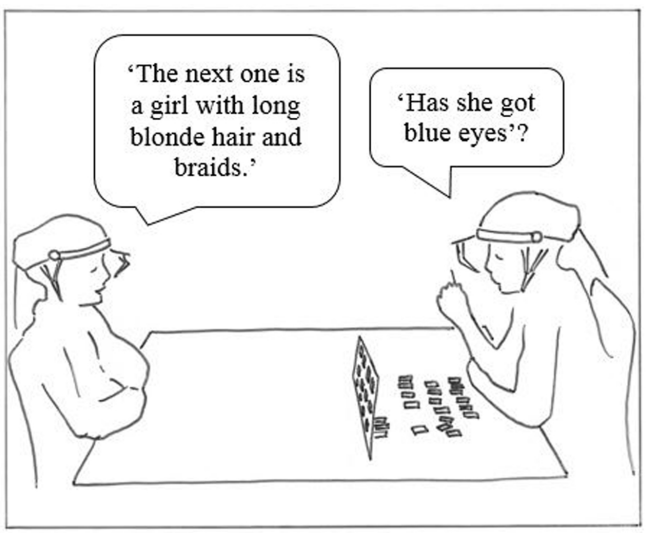


FIGURE 1 | Sketch of the experimental setting showing the speaker (on the left) describing pictures of faces, and the listener (on the right) requesting additional information. Adapted from Sandgren et al. (2012).

Referential communication requires a basic level of linguistic skills but also a range of cognitive capacities. The linguistic information must be processed and maintained until a referent has been chosen, requiring working memory capacity (WMC), the demands on which are likely to vary depending on the description provided (Dahlgren and Dahlgren Sandberg, 2008). Finally, in order for the speaker to provide an adequately detailed description, and for the listener to adjust his/her questions appropriately, both interlocutors must be able to take the perspective of the conversational partner.

REFERENTIAL COMMUNICATION—FINDINGS

In a range of studies we have used an adapted version of the referential communication task, as a complement to linguistic and cognitive assessment, to investigate the communicative abilities of Swedish speaking children and adolescents with varying degrees of SNHI. While conducting the experiments under optimal acoustic conditions, with rigid experimental control, participants were instructed to choose a friend with whom to complete the task, thereby maintaining ecological validity. In the first study, Ibertsson et al. (2009a) found 11 to 19 year-old adolescents with severe-to-profound SNHI and cochlear implants to request more information than NH peers to resolve ambiguities caused by inaccurate or insufficient information from the conversational partner. The participants showed an increased use of requests for confirmation (yes/no questions, for example, "Does she have blonde hair?"), as compared to requests for elaboration ("What color is her hair?"). This use of questions was interpreted as a conversational strategy aimed at limiting the number of possible responses from the partner and thereby reducing the risk of misunderstanding. This conversational strategy was found to be related to complex WMC (Ibertsson et al., 2009b). Participants with SNHI and reduced WMC were found to use requests for confirmation of information mentioned earlier in the conversation ("Did you say he had a beard?") whereas participants with greater WMC requested confirmation of new information to a greater extent, more clearly driving the conversation forward (Ibertsson et al., 2009b). Responses to the requests have not been shown to differ between the groups (Sandgren et al., 2011).

In an effort to obtain a fuller picture of the communicative exchanges during referential communication—both speech and body communication—we recently fitted interlocutors with mobile eye trackers (Sandgren et al., 2012, 2013, 2014). We were able to show that moments of mutual gaze, in which the listener looks at the speaker, showed a tight temporal connection with important parts of the spoken message (Sandgren et al., 2012). Questions, back-channeling responses, and statements, directed

from the listener to the speaker, were all associated with higher probability of listener gaze to the speaker's face. The results indicate that the spoken message is emphasized by the gaze exchanges, even to the point of making the content of the spoken message relevant. In a recurring example from the data, questions remained unanswered when not accompanied by a gaze to the respondent's face (Sandgren et al., 2012). In a comparison between 10 and 15 year-old children with mild-to-moderate SNHI (mean age 12;6, SD 2;0; mean better ear pure-tone average 33.0 dB HL, SD 7.8) and NH same-age peers, the gaze behavior was found to be accentuated in the participants with SNHI, showing greater odds (ORs 1.2–2.1) for gaze to the speaker's face than NH peers (Sandgren et al., 2014).

Since other factors than hearing differ between children with and without hearing impairment, we went on to investigate group differences in the probability of gaze to the speaker's face while adjusting for individual performance on receptive grammar, expressive vocabulary, complex WMC, and phonological short term memory (PSTM; Sandgren et al., 2013). In the collected sample (cf. Sandgren et al., 2014 above), children with SNHI performed significantly below NH controls on non-word repetition (measuring phonological processing and PSTM) and expressive vocabulary, while non-significant differences were found for receptive grammar and complex WMC.

The group difference in gaze behavior remained significant despite adjustment for receptive grammar, expressive vocabulary, and complex WMC, but not non-word repetition, revealing an interaction between SNHI and PSTM capacity. Participants with SNHI with lower scores on non-word repetition (>1.25 SD below NH mean) showed a twofold increase in the probability of gaze to the speaker's face, whereas those with higher scores had a reduced probability of looking at the conversational partner (Sandgren et al., 2013).

CONCLUSIONS AND IMPLICATIONS

To summarize, request strategies and gaze behavior in children with SNHI during referential communication represent control and validation mechanisms which go above and beyond what is explained by the hearing impairment alone, and the results highlight WMC and PSTM capacity as driving forces behind the effect. While active and competent conversational partners, the participants with SNHI exhibit conversational strategies distinct from those of NH peers despite optimal conditions (clear task objectives, known conversational partner, no time limit, and silent surroundings). The findings affect clinical and school-based management of hearing impairment as well as our theoretical assumptions of the course of development of hearing impairment and its consequences. Speech-language pathologists, audiologists, psychologists and teachers working with children with SNHI should be aware of an increased likelihood of language deficits, which require intervention and adaptations to ensure academic attainment. This is equally relevant for younger school-aged children, whose language deficits may be easy to detect, and for later school years, when language profiles may have changed and previously sufficient coping strategies are challenged as school demands increase and learning is expected in adverse listening conditions (Bishop, 2014). Relevant for all is a comprehensive and continual evaluation of communicative functioning, including formal assessment of language, cognition, and interaction.

Our findings support the notion of WMC and PSTM playing important roles in the integration of auditory and visual information during speech production and perception. As suggested by the Ease of Language Understanding model (Rönnberg et al., 2013), a mismatch between input and long term memory representations will evoke extrinsic processing of the acoustic signal, requiring cognitive effort and strategic use of multimodal information, in this case possibly increased use of questions and gaze behavior during conversation. Future studies should evaluate individual variability in these memory capacities in relation to contextual multimodal challenge and support in the search for an explanation for the heterogeneity in language and communication outcome for children with SNHI. This should also provide an answer to whether the changes in request strategies and gaze behavior are, indeed, compensatory. The need for thorough and systematic studies of communication in children with SNHI should, however, not preclude prompt implementation of effective interventions based on current theories of language learning in typical and atypical populations.

REFERENCES

Asker-Árnason, L., Ibertsson, T., Wass, M., Wengelin, Å., and Sahlén, B. (2010). Picture-elicited written narratives, process and product, in 18 children with cochlear implants. *Commun. Disord. Q.* 31, 195–212. doi: 10.1177/1525740109337734

Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *Int. J. Lang. Commun. Disord.* 49, 381–415. doi: 10.1111/1460-6984.12101

Briscoe, J., Bishop, D. V. M., and Norbury, C. F. (2001). Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *J. Child Psychol. Psychiatry* 42, 329. doi: 10.1111/1469-7610.00726

Dahlgren, S., and Dahlgren Sandberg, A. (2008). Referential communication in children with autism spectrum disorder. *Autism* 12, 335–348. doi: 10.1177/1362361308091648

Gilbertson, M., and Kamhi, A. G. (1995). Novel word learning in children with hearing impairment. *J. Speech Hear. Res.* 38, 630–642. doi: 10.1044/jshr.3803.630

Glucksberg, S., and Krauss, R. M. (1967). What do people say after they have learned to talk? Studies of the development of referential communication. *Merrill Palmer Q. Behav. Dev.* 13, 309–316.

Hansson, K., Forsberg, J., Löfqvist, A., Mäki-Torkko, E., and Sahlén, B. (2004). Working memory and novel word learning in children with hearing impairment and children with specific language impairment. *Int. J. Lang. Commun. Disord.* 39, 401–422. doi: 10.1080/13682820410001669887

Hansson, K., Sahlén, B., and Mäki-Torkko, E. (2007). Can a 'single hit' cause limitations in language development? A comparative study of Swedish children with hearing impairment and children with specific language impairment. *Int. J. Lang. Commun. Disord.* 42, 307–323. doi: 10.1080/13682820600933526

Ibertsson, T., Hansson, K., Mäki-Torkko, E., Willstedt-Svensson, U., and Sahlén, B. (2009a). Deaf teenagers with cochlear implants in conversation with hearing peers. *Int. J. Lang. Commun. Disord.* 44, 319–337. doi: 10.1080/13682820802052067

Ibertsson, T., Hansson, K., Asker-Arnason, L., and Sahlén, B. (2009b). Speech recognition, working memory and conversation in children with cochlear implants. *Deafness Educ. Int.* 11, 132–151. doi: 10.1179/146431509790559615

Mayne, A. M., Yoshinaga-Itano, C., and Sedey, A. L. (1998a). Receptive vocabulary development of infants and toddlers who are deaf or hard of hearing. *Volta Rev.* 100, 29–52.

Mayne, A. M., Yoshinaga-Itano, C., Sedey, A. L., and Carey, A. (1998b). Expressive vocabulary development of infants and toddlers who are deaf or hard of hearing. *Volta Rev.* 100, 1–28.

Most, T., Shina-August, E., and Meilijson, S. (2010). Pragmatic abilities of children with hearing loss using cochlear implants or hearing AIDS compared to hearing children. *J. Deaf Stud. Deaf Educ.* 15, 422–437. doi: 10.1093/deafed/enq032

Norbury, C. F., Bishop, D. V. M., and Briscoe, J. (2001). Production of English finite verb morphology: a comparison of SLI and mild-moderate hearing impairment. *J. Speech Lang. Hear. Res.* 44, 165–178. doi: 10.1044/1092-4388(2001/015)

Perkins, M. (2007). *Pragmatic Impairment*. Cambridge: Cambridge University Press.

Reuterskiöld, C., Ibertsson, T., and Sahlén, B. (2010). Venturing beyond the sentence level: narrative skills in children with hearing loss. *Volta Rev.* 110, 389–406.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Sahlén, B., and Hansson, K. (2006). Novel word learning and its relation to working memory and language in children with mild-to-moderate hearing impairment and children with specific language impairment. *J. Multilin. Commun. Disord.* 4, 95–107. doi: 10.1080/14769670600929360

Sahlén, B., Hansson, K., Ibertsson, T., and Reuterskiöld Wagner, C. (2004). Reading in children of primary school age—a comparative study of children with hearing impairment and children with Specific Language Impairment. *Acta Neuropsychol.* 2, 393–407.

Sandgren, O., Andersson, R., Van De Weijer, J., Hansson, K., and Sahlén, B. (2012). Timing of gazes in child dialogues: a time-course analysis of requests and back channelling in referential communication. *Int. J. Lang. Commun. Disord.* 47, 373–383. doi: 10.1111/j.1460-6984.2012.00151.x

Sandgren, O., Andersson, R., Van De Weijer, J., Hansson, K., and Sahlén, B. (2013). Impact of cognitive and linguistic ability on gaze behavior in children with hearing impairment. *Front. Psychol.* 4:856. doi: 10.3389/fpsyg.2013.00856

Sandgren, O., Andersson, R., Van De Weijer, J., Hansson, K., and Sahlén, B. (2014). Coordination of gaze and speech in communication between children with hearing impairment and normal-hearing peers. *J. Speech Lang. Hear. Res.* 57, 942–951. doi: 10.1044/2013_JSLHR-L-12-0333

Sandgren, O., Ibertsson, T., Andersson, R., Hansson, K., and Sahlén, B. (2011). 'You sometimes get more than you ask for': responses in referential communication between children and adolescents with cochlear implant and hearing peers. *Int. J. Lang. Commun. Disord.* 46, 375–385. doi: 10.3109/13682822.2010.507617

Toe, D. M., and Paatsch, L. E. (2010). The communication skills used by deaf children and their hearing peers in a question-and-answer game context. *J. Deaf Stud. Deaf Educ.* 15, 228–241. doi: 10.1093/deafed/enq006

Wake, M., Tobin, S., Cone-Wesson, B., Dahl, H.-H., Gillam, L., Mccormick, L., et al. (2006). Slight/mild sensorineural hearing loss in children. *Pediatrics* 118, 1842–1851. doi: 10.1542/peds.2005-3168

Wass, M., Ibertsson, T., Lyxell, B., Sahlén, B., Hällgren, M., Larsby, B., and Mäki-Torkko, E. (2008). Cognitive and linguistic skills in Swedish children with cochlear implants—measures of accuracy and latency as indicators of development. *Scand. J. Psychol.* 49, 559–576. doi: 10.1111/j.1467-9450.2008.00680.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 January 2015; paper pending published: 06 February 2015; accepted: 17 February 2015; published online: 09 March 2015.

Citation: Sandgren O, Hansson K and Sahlén B (2015) Working memory and referential communication—multimodal aspects of interaction between children with sensorineural hearing impairment and normal hearing peers. Front. Psychol. ***6****:242. doi: 10.3389/fpsyg.2015.00242*

This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology.

Copyright © 2015 Sandgren, Hansson and Sahlén. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

PERSPECTIVE
published: 17 July 2015
doi: 10.3389/fpsyg.2015.00998

Oral communication in individuals with hearing impairment—considerations regarding attentional, cognitive and social resources

Ulrike Lemke[*†] and Sigrid Scherpiet*[†]

Cognitive and Ecological Audiology, Science and Technology, Phonak AG, Stäfa, Switzerland

Traditionally, audiology research has focused primarily on hearing and related disorders. In recent years, however, growing interest and insight has developed into the interaction of hearing and cognition. This applies to a person's listening and speech comprehension ability and the neural realization thereof. The present perspective extends this view to oral communication, when two or more people interact in social context. Specifically, the impact of hearing impairment and cognitive changes with age is discussed. In focus are executive functions, a group of top-down processes that guide attention, thought and action according to goals and intentions. The strategic allocation of the limited cognitive processing capacity among concurrent tasks is often effortful, especially under adverse communication conditions and in old age. Working memory, a sub-function extensively discussed in cognitive hearing science, is here put into the context of other executive and cognitive functions required for oral communication and speech comprehension. Finally, taking an ecological view on hearing impairment, activity limitations and participation restrictions are discussed regarding their psycho-social impact and third-party disability.

Keywords: communication, hearing impairment, executive functions, cognitive aging, speech comprehension, third-party disability

OPEN ACCESS

Edited by:
Mary Rudner, Linköping University, Sweden

Reviewed by:
Melanie A. Ferguson, Nottingham University Hospitals NHS Trust, UK
Stig D. Arlinger, Linköping University, Sweden

***Correspondence:**
Ulrike Lemke, Cognitive and Ecological Audiology, Science and Technology, Phonak AG, Laubisrütistrasse 28, CH-8712 Stäfa, Switzerland
ulrike.lemke@phonak.com

[†]*These authors have contributed equally to this work.*

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *05 March 2015*
Accepted: *02 July 2015*
Published: *17 July 2015*

Citation:
Lemke U and Scherpiet S (2015) Oral communication in individuals with hearing impairment—considerations regarding attentional, cognitive and social resources. Front. Psychol. 6:998. doi: 10.3389/fpsyg.2015.00998

General Aspects of Oral Communication

Being able to communicate with others is regarded a key element of human functioning. During oral communication individuals interact with each other, and also with their social and physical surroundings by exchanging information in form of language, signals, and behavior (Stephens and Kramer, 2009). As such oral communication constitutes by far a more complex process than serving the basic purpose of sending and receiving information. Communication implies bidirectional transfer of information, meaning, and intent between two or more individuals (Kiessling et al., 2003). As such, it is a social act originating from the need to express oneself, and to relate to others. Furthermore, interactions are mediated by psychological variables of the communication partners such as emotions, attitudes, and beliefs as well as by values and rules of the community. Thus, oral communication is a broad concept encompassing perceptual, cognitive, psychological, and social constructs.

Hearing impairment constitutes a major challenge in this respect as it generally leads to difficulties in oral communication (Stephens and Jones, 2005). These communication problems are often age-related and accompanied by impairment of other sensory modalities and comorbid health problems (Kramer et al., 2002; Davis and Davis, 2009; Lemke, 2009; Stam et al., 2014). Age-related hearing

impairment (presbycusis) begins in the fourth decade and its prevalence increases with age. About half of the population over the age of 65 years and up to 90% of individuals over the age of 80 years are affected by presbycusis (Cruikshanks et al., 2010; Lin et al., 2011c). The consequences of hearing impairment can be far reaching, commonly affecting not only the hearing impaired person, but also their communication partners, primarily significant others (SOs), and social networks. According to the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) (WHO; World Health Organization, 2001) communication disability due to hearing impairment is an outcome of interactions between sensory impairment and participation in life. For instance, hearing impairment often makes it difficult to participate in social and cultural activities due to a restricted ability to interact and communicate with peers. This can lead to withdrawal from activities and participation potentially resulting in feelings of loneliness and social isolation (Pronk et al., 2011).

As auditory perception sets the basis for oral communication, the contribution of the auditory system often is narrowed down to the term "hearing." However, the concept should be disentangled and extended into more specific mechanisms that drive the stream of oral communication, that is *hearing*, *listening*, *comprehending*, and eventually *communicating* (World Health Organization, 2001; Kiessling et al., 2003). In the communication pathway *hearing* represents an important, rather passive function denoting the perception of sound. It is usually at this stage of sensory processing that hearing impairment is described by means of audiometry. *Listening*, *comprehending*, and *communicating* on the other hand are considered more complex processes that require active engagement of the individual(s) as well as fast interactions between sensory and cognitive processing. For example, *listening* to someone can be referred to as hearing with intention and attention. As such listening often demands the expenditure of mental effort, because cognitive resources including attention and executive functions (EFs) have to be invested for goal pursuit. Besides, the information must be received and decoded in a unidirectional manner in order to be able to derive and understand meaning. This step is described as *comprehension* and takes place throughout conversations with others. Finally, *communication* involves the conversational interactions between two or more people, while transferring information, meaning and intent bi-directionally. Given the described steps in the communication pathway, successful oral communication depends not only on the ability of hearing, but also requires listening and comprehending from all participants involved. One could understand a communication situation as a dynamic system that must be carefully balanced. Difficulties in either one component, such as one communication partner being hearing impaired, would require sensitivity and flexibility by means of adaptation of the system. To maintain the flow of a conversation and to avoid interruptions, when communication problems occur, strategies for compensation and repair need be activated immediately. Such strategies could include that the speaker repeats or rephrases what was said using loud and clear voicing, or that the hearing impaired person tries to concentrate more and activates additional mental resources (e.g., filling the gaps through context) or relies more on other modalities (e.g., visual cues for lip-reading; Lind, 2009; Lind et al., 2010).

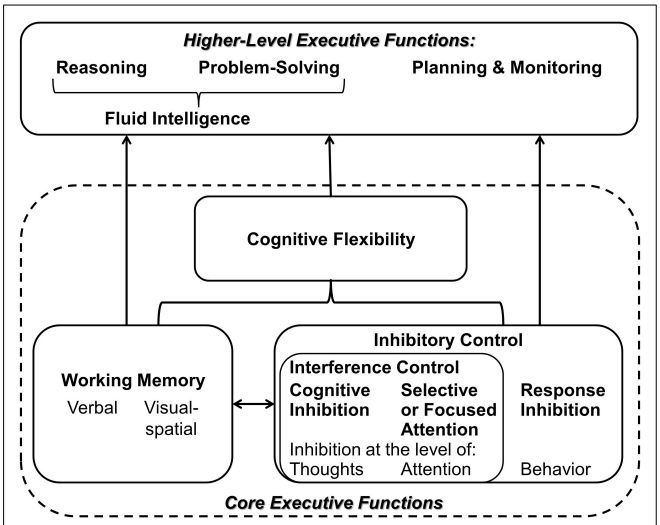


FIGURE 1 | Executive functions—a set of top-down mental processes (adapted from Diamond, 2013).

Executive Functions and Attention Steer Oral Communication

Oral communication requires concentration and paying attention, thus demanding specific top-down mental processes that are referred to as EFs (Miller and Cohen, 2001; Burgess and Simons, 2005; Diamond, 2013). These EFs enable the strategic handling of communicational intentions such as taking time to think before responding, considering unanticipated arguments, resisting the temptation to interrupt a communication partner, and staying focused throughout a conversation. **Figure 1** shows EFs that have consistently been identified and that have been associated with a prefrontal-parietal neural network (Diamond, 2013). While there is inconsistency in the literature regarding the use of specific terms and the modeling of EFs, there is general agreement on three essential functions behind this network, namely *inhibitory and interference control*, *working memory*, and *cognitive flexibility* (Miyake et al., 2000; Miller and Cohen, 2001; Diamond, 2013). These core functions mediate higher order EFs such as reasoning, problem solving (the latter two being used synonymous with fluid intelligence), anticipation and planning. Overall, EFs describe the ability to guide attention, thought, and action in accord with goals or intentions as it is required in oral communication (Miller and Cohen, 2001).

The degree to which attention, EF and other cognitive resources have to be allocated and engaged for a specific listening goal is referred to as listening effort, which is especially reported under adverse listening conditions and for cognitively demanding listening tasks (Anderson Gosselin and Gagne, 2011; Picou et al., 2011; McGarrigle et al., 2014; and respective comments from Ronnberg et al., 2014; Wingfield, 2014). This is for instance the case, when auditory perception is compromised by distracting background noise, reverberant conditions, competing voices,

and/or a degraded auditory signal due to a person's hearing impairment (Arlinger et al., 2009). Under such circumstances, there is a high demand for core EFs, which is especially challenging in old age and will be outlined in more detail below (Erb and Obleser, 2013).

Firstly, *inhibitory and interference control* enable the selective allocation and reallocation of attention. Thus, it becomes for instance possible to focus on the voice of interest in a multi-talker environment, while suppressing other auditory streams. In hearing impairment, degraded signals trigger automatic, stimulus driven, bottom-up processing. Because they are more difficult to analyze they attract additional involuntary attention (Shinn-Cunningham, 2007; Shinn-Cunningham and Best, 2008). Consequently, it becomes more demanding to ignore or attend to specific stimuli driven by top-down goals and intentions (Posner and DiGirolamo, 1998). Also, it should be noted that older adults tend to develop difficulties in inhibition of distractions (Alain and Woods, 1999). While the ability of focusing attention usually remains intact in old age, there is strong evidence for an inhibitory-control deficit in aging (Gazzaley et al., 2005; Diamond, 2013). This age-specific difficulty is most probably taking its toll in complex communication situations and to an even greater extent in the presence of hearing impairment.

Secondly, the core EF of *working memory* (WM)—the ability to hold information in mind (maintain) and mentally work with it (manipulate) at the same time (Baddeley, 1992)—is key for speech understanding and communication (cf. new ELU-model; Ronnberg et al., 2013). WM allows one to relate things to each other over time, to consider alternatives, and to make decisions considering the past and the future. With regard to WM, evidence is in support of models that suggest a functional (maintenance vs. manipulation) as well as domain-specific organization (verbal vs. visual-spatial) in the frontal brain (Ullsperger and von Cramon, 2006). In the context of oral communication, verbal WM is necessary for comprehending speech, when meaning unfolds over the course of words and sentences. Nonetheless, visual-spatial WM can also play a role in the analysis of an auditory communication scene, as it facilitates the localization and segregation of speakers and other audio sources. Hearing impairment additionally loads on WM (e.g., when degraded information has to be put in context to derive its meaning; Ronnberg et al., 2013). Also here, a decline in WM capacity is common with age (e.g., Park, 1999) and constitutes an additional challenge for individuals with age-related hearing impairment. To a great extent this decline in WM seems to be due to the decline in inhibitory control (Hedden and Park, 2001). Moreover, a big overlap of age-related changes in speed of information processing and WM has been observed and controversially discussed (Salthouse, 1992; Zimprich and Kurtz, 2013). Inhibitory control and WM support each other. For example, in order to follow and participate in a conversation distracting thoughts and lines have to be disregarded and relevant information has to be retained.

Thirdly, it should be noted that the two previously discussed functions together provide the basis for a third core EF, which is *cognitive flexibility*. It describes the ability to change perspective regarding a problem, to be creative, to adjust to new demands, or to switch tasks according to priorities (Diamond, 2013). In general, cognitive flexibility also declines with age. For instance, in tasks that require switching between rules or response sets, older adults tend to slow down to maintain accuracy (Kray and Lindenberger, 2000; Cepeda et al., 2001). Older adults tend to recruit EFs in a rather reactive way in response to demands, whereas young adults tend to be anticipatory and proactively in recruiting EFs (Karayanidis et al., 2011).

Cognitive Resources for Speech Comprehension

Central to oral communication is the ability to understand speech, which entails constant interactions between auditory and cognitive processing (Pichora-Fuller and Singh, 2006). Sounds continuously arrive at the ears via vibrations of air and are converted to linguistic representations in the brain (Craik, 2007). It is a bidirectional process taking in bottom-up information by using the perceptual system and conveying these inputs with top-down knowledge that has developed through experience (Pichora-Fuller, 2008a). Good quality of the signal facilitates speech understanding and better cognitive resources increase the chances to understand. In more detail, the bottom-up perspective is referred to as data-driven processing that involves mechanisms of conveying information from acoustic signals to phonemes, words, phrases and sentences. It is based on peripheral auditory processes that depend on the perceptual accuracy in coding and transferring acoustic information. Top-down effects, on the other side, are conceptually-driven cognitive processes that enable speech perception by linguistic context and expectation of the listener using the influence of memories and knowledge (Norman and Bobrow, 1975). Cognitive domains that apply for successful speech understanding primarily include speed of information processing (Review: Schneider et al., 2010), selective focused attention (e.g., Koelewijn et al., 2014), WM (e.g., Baddeley, 1992; Akeroyd, 2008) as well as semantic knowledge, namely language abilities and context integration (e.g., Pichora-Fuller and Singh, 2006; Zekveld et al., 2011).

In normal hearing individuals, the abilities to segregate, select, store, identify, and integrate information is often at risk in complex or adverse listening conditions. In case of hearing impairment and/or old age, additional challenges are introduced by compromised bottom-up information and/or decrements in top-down cognitive resources due to age-related changes (e.g., Bregman, 1990; Pichora-Fuller, 2003). Cognitive resources are generally limited and their processing "capacity" is assigned and flexibly shared between a number of tasks according to priorities (Moray, 1967; Kahneman, 1973; Wickens, 2008). In order to compensate for auditory deficits, hearing impaired listeners must invest more cognitive resources, for instance in order to follow a conversation. This is typically perceived as effortful by the listener. Also, these resources might otherwise be available for parallel tasks. In demanding listening situations, cognitive resources, such as rapidly switching attention and suppressing interfering sounds are additionally needed to extract the speech signal from competing sound

sources and then to match it to mental representations of the phonological and semantic long term memory (cf. new ELU-model; Ronnberg et al., 2013). Consequently, less cognitive capacity is reserved and available for additional processes such as maintenance and manipulation of novel auditory information in WM or establishing episodic memory traces (Tun and Wingfield, 1999; Wingfield and Tun, 2001; Rudner et al., 2011; Mishra, 2014; Rudner and Lunner, 2014). In other words, speech understanding under adverse conditions takes up more cognitive capacity, firstly to decode the speech signal and secondly to comprehend it in order to be able to communicate, respectively. At this level, typical age-relevant cognitive declines in speed of information processing, inhibitory and interference control, WM capacity, and/or mental flexibility described earlier, may contribute even more to communication difficulties for the listener. Nevertheless, some of the above mentioned age-related challenges might be compensated if context information becomes available. For example, it has been shown that older people have a broader semantic knowledge and vocabulary, wider social experiences in a variety of communication situations, and make better use of prosody and context compared to younger individuals (Pichora-Fuller and Singh, 2006; Pichora-Fuller, 2008b).

Overall, speech understanding is realized through a widespread neural circuit that is mapped as a dynamic temporo-frontal network in the brain. Bottom-up information arrives at the auditory cortex within the temporal lobe and is directed to higher-order brain regions of the frontal cortex along multiple long-range language connections specified by ventral and dorsal pathways (Friederici, 2012). The ventral pathway is associated with the processing of sound-to-meaning and has been suggested to map acoustic speech signals onto lexical conceptual representations. The dorsal stream, on the other hand, is linked to the processing of sound-to-action and has been proposed the role of mapping signals onto articulatory motor representations (Hickok and Poeppel, 2007). More specifically, language-related brain areas typically comprise Broca's area in the inferior frontal gyrus, Wernicke's area in the superior temporal gyrus, and also parts of the middle temporal gyrus and the inferior parietal regions (Friederici, 2011). In this respect, the temporal cortex plays an important role for oral communication, given that this is the center where further connections for higher order processing are linked enabling the integration of attention, memory, and context for understanding speech. Interestingly enough, brain imaging studies have shown that with increasing age physiological changes in the healthy brain may become relevant for the integration of different cognitive resources in speech understanding under challenging listening conditions. These changes include reduced connectivity of neurons and thus interactions between brain regions; moderate loss of brain mass especially in the prefrontal cortex, medial temporal cortex (esp. hippocampus) or caudate nucleus; as well as changes in neurotransmitter systems such as the dopaminergic systems (e.g., Raz, 2005; Park and Reuter-Lorenz, 2009). Yet, literature has also shown that compensatory effects in old age as mentioned earlier are also reflected by brain activation patterns. A more extensive brain activity has been observed when listeners engaged in additional top-down context-driven processing (Davis et al., 2005). Primarily, activations in areas of the prefrontal and parietal cortices during listening in adverse conditions suggest increased functional connectivity between high-order cortical areas and indicate the allocation of additional, especially executive resources for semantic processing (Obleser et al., 2007). These widespread activations support compensatory processing in old age (Cabeza et al., 2002).

As there is a close association of hearing impairment and cognitive decline in old age, several explanations have been proposed and are debated (Li and Lindenberger, 2002; Lin et al., 2011a,d). Importantly, none of the explanatory models are exclusive, but instead could be coexistent. Namely, it has been hypothesized that sensory and cognitive decline in old age share their pathologic etiology and have a "common cause." Also, the described interaction of hearing impairment and cognitive load in the sense of resource competition and limited capacity could explain this association. Last but not least, social and psychological factors have to be taken into account as the interaction of hearing and cognition could be mediated through those.

Social Resources and Consequences

Considering that communication takes place between two or more individuals and in the context of culture and society, it is influenced by shared and unshared patterns of action, meaning, and values. These phenomena are intensively studied in social psychology with regard to intrapersonal (e.g., self-concept and social cognition) and interpersonal processes (e.g., social influence, group dynamics, attractions, and generation gap; e.g., Tesser and Schwarz, 2000; Fletcher and Clark, 2002). One's thoughts, feelings, and behaviors are influenced by the presence of others and interaction with others. Therefore, difficulties in communication that are driven by hearing impairment may have significant consequences concerning the sense of security in everyday life, quality of life, social and emotional functioning as well as psychological wellbeing (e.g., Strawbridge et al., 2000; Nachtegaal et al., 2009, 2012; Danermark et al., 2010; Stam et al., 2013, 2014; Hogan et al., 2015). It is evident that poor hearing leads to communication impairments that may result in social isolation and may mediate disadvantageous health and functional consequences (Berkman et al., 2000; Uchino, 2006). Also, it is hypothesized that withdrawal from social participation may put hearing impaired people at risk for more rapid cognitive decline (Uhlmann et al., 1989; Gates et al., 2010, 2011; Lin et al., 2011a,b). In connection with this, the role of communication partners, especially SOs, in hearing impairment has gained interest during recent years. This is particularly evident in the WHO's ICF classification (World Health Organization, 2001) of the effect of hearing impairment on SOs as a third-party disability. Third-party disability is described to occur when the SO does not have a hearing impairment themselves, but experiences activity limitations and participation restrictions as a result of their partner's hearing impairment (Scarinci et al., 2009, 2012). SOs are reported to experience a restricted social life, increased burden of communication, and poorer quality of life and relationship

satisfaction (Kamil and Lin, 2015). Treatment of hearing impairment that typically comprises hearing aids, cochlear implants, and audiological rehabilitation programs targeting the hearing impaired person, tend to also improve quality of life, communication, feelings toward the hearing impaired person, and activity participation of the SO (Kamil and Lin, 2015).

Concluding Remarks

Modern audiology has extended its focus from hearing to considerations of cognitive processes, aging effects and social factors in order to address the communication problems of hearing impaired individuals and to meet their expectations. In recent years, great insight has been gained into this interdisciplinary field of study. For instance research has taken into account aspects of neuro-cognitive mechanisms, age-related decrements and compensatory strategies, as well as the role of SOs and the social network related to successful oral communication and rehabilitation. Nevertheless, there is still great potential for applying this knowledge as a matter of course in aural rehabilitation (e.g.,Ekberg et al., 2014, 2015; Hickson et al., 2014) and translating it into services and products to the benefit of the hearing impaired (e.g., Lunner et al., 2009; Pichora-Fuller et al., 2013).

Acknowledgments

Both authors are employed at Phonak AG, Switzerland. The present work was conducted in Phonak's research program Cognitive and Ecological Audiology at the department of Science and Technology.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

Alain, C., and Woods, D. L. (1999). Age-related changes in processing auditory stimuli during visual attention: evidence for deficits in inhibitory control and sensory memory. *Psychol. Aging* 14, 507–519. doi: 10.1037/0882-7974.14.3.507

Anderson Gosselin, P., and Gagne, J. P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise. *J. Speech Lang. Hear. Res.* 54, 944–958. doi: 10.1044/1092-4388(2010/10-0069)

Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x

Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359

Berkman, L. F., Glass, T., Brissette, I., and Seeman, T. E. (2000). From social integration to health: durkheim in the new millennium. *Soc. Sci. Med.* 51, 843–857. doi: 10.1016/S0277-9536(00)00065-4

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Burgess, P. W., and Simons, J. S. (2005). "Theories of frontal lobe executive function: clinical applications," in *Effectiveness of Rehabilitation for Cognitive Deficits*, eds P. W. Halligan and D. T. Wade (Oxford: Oxford University Press), 211–231.

Cabeza, R., Anderson, N. D., Locantore, J. K., and Mcintosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402. doi: 10.1006/nimg.2002.1280

Cepeda, N. J., Kramer, A. F., and Gonzalez De Sather, J. C. (2001). Changes in executive control across the life span: examination of task-switching performance. *Dev. Psychol.* 37, 715–730. doi: 10.1037/0012-1649.37.5.715

Craik, F. I. (2007). The role of cognition in age-related hearing loss. *J. Am. Acad. Audiol.* 18, 539–547. doi: 10.3766/jaaa.18.7.2

Cruikshanks, K. J., Zhan, W., and Zhong, W. (2010). "Epidemiology of age-related hearing impairment," in *The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis*, eds S. Gordon-Salant, R. D. Frisina, A. Pop-per, and D Fay (Berlin: Springer), 259–274.

Danermark, B., Cieza, A., Gange, J. P., Gimigliano, F., Granberg, S., Hickson, L., et al. (2010). International classification of functioning, disability, and health core sets for hearing loss: a discussion paper and invitation. *Int. J. Audiol.* 49, 256–262. doi: 10.3109/14992020903410110

Davis, A., and Davis, K. A. (2009). "Epidemiology of aging and hearing loss related to other chronic illnesses," in *Hearing Care for Adults 2009—The Challenge of Aging. Proceedings of the 2nd International Adult Conference*, ed. L. Hickson (Staefa: Phonak AG), 23–32.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., and Mcgettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *J. Exp. Psychol. Gen.* 134, 222–241. doi: 10.1037/0096-3445.134.2.222

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Ekberg, K. C. M., Scarinci, N., Grenness, C., and Hickson, L. (2014). Disagreements between clients and family members regarding clients' hearing and rehabilitation within audiology appointments for older adults. *J. Interact. Res. Commun. Disord.* 5, 217–244. doi: 10.1558/jircd.v5i2.217

Ekberg, K., Meyer, C., Scarinci, N., Grenness, C., and Hickson, L. (2015). Family member involvement in audiology appointments with older people with hearing impairment. *Int. J. Audiol.* 54, 70–76. doi: 10.3109/14992027.2014.948218

Erb, J., and Obleser, J. (2013). Upregulation of cognitive control networks in older adults' speech comprehension. *Front. Syst. Neurosci.* 7:116. doi: 10.3389/fnsys.2013.00116

Fletcher, G., and Clark, M. S. (2002). *Blackwell Handbook of Social Psychology: Interpersonal Processes.* Oxford: Wiley-Blackwell.

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn. Sci.* 16, 262–268. doi: 10.1016/j.tics.2012.04.001

Gates, G. A., Anderson, M. L., Mccurry, S. M., Feeney, M. P., and Larson, E. B. (2011). Central auditory dysfunction as a harbinger of alzheimer dementia. *Arch. Otolaryngol. Head Neck Surg.* 137, 390–395. doi: 10.1001/archoto.2011.28

Gates, G. A., Gibbons, L. E., Mccurry, S. M., Crane, P. K., Feeney, M. P., and Larson, E. B. (2010). Executive dysfunction and presbycusis in older persons with and without memory loss and dementia. *Cogn. Behav. Neurol.* 23, 218–223. doi: 10.1097/WNN.0b013e3181d748d7

Gazzaley, A., Cooney, J. W., Mcevoy, K., Knight, R. T., and D'esposito, M. (2005). Top-down enhancement and suppression of the magnitude and speed of neural activity. *J. Cogn. Neurosci.* 17, 507–517. doi: 10.1162/0898929053279522

Hedden, T., and Park, D. (2001). Aging and interference in verbal working memory. *Psychol. Aging* 16, 666–681. doi: 10.1037/0882-7974.16.4.666

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hickson, L., Meyer, C., Lovelock, K., Lampert, M., and Khan, A. (2014). Factors associated with success with hearing aids in older adults. *Int. J. Audiol.* 53(Suppl. 1), S18–S27. doi: 10.3109/14992027.2013.860488

Hogan, A., Phillips, R. L., Brumby, S. A., Williams, W., and Mercer-Grant, C. (2015). Higher social distress and lower psycho-social wellbeing: examining the coping capacity and health of people with hearing impairment. *Disabil. Rehabil.* 6, 1–6. doi: 10.3109/09638288.2014.996675

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.

Kamil, R. J., and Lin, F. R. (2015). The effects of hearing impairment in older adults on communication partners: a systematic review. *J. Am. Acad. Audiol.* 26, 155–182. doi: 10.3766/jaaa.26.2.6

Karayanidis, F., Whitson, L. R., Heathcote, A., and Michie, P. T. (2011). Variability in proactive and reactive cognitive control processes across the adult lifespan. *Front. Psychol.* 2:318. doi: 10.3389/fpsyg.2011.00318

Kiessling, J., Pichora-Fuller, M. K., Gatehouse, S., Stephens, D., Arlinger, S., Chisolm, T., et al. (2003). Candidature for and delivery of audiological services: special needs of older people. *Int. J. Audiol.* 42(Suppl. 2), S92–S101. doi: 10.3109/14992020309074650

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., and Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. *Hear. Res.* 312, 114–120. doi: 10.1016/j.heares.2014.03.010

Kramer, S. E., Kapteyn, T. S., Kuik, D. J., and Deeg, D. J. (2002). The association of hearing impairment and chronic diseases with psychosocial health status in older age. *J Aging Health* 14, 122–137. doi: 10.1177/089826430201400107

Kray, J., and Lindenberger, U. (2000). Adult age differences in task switching. *Psychol. Aging* 15, 126–147. doi: 10.1037/0882-7974.15.1.126

Lemke, U. (2009). "The challenges of aging—Sensory, cognitive, socio-emotional, and health changes in old age," in *Hearing Care for Adults 200—The Challenge of Aging. Proceedings of the 2nd International Adult Conference*, ed. L. Hickson (Staefa: Phonak AG), 33–43.

Li, K. Z., and Lindenberger, U. (2002). Relations between aging sensory/sensorimotor and cognitive functions. *Neurosci. Biobehav. Rev.* 26, 777–783. doi: 10.1016/S0149-7634(02)00073-8

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., and Resnick, S. M. (2011a). Hearing loss and cognition in the baltimore longitudinal study of aging. *Neuropsychology* 25, 763–770. doi: 10.1037/a0024238

Lin, F. R., Metter, E. J., O'brien, R. J., Resnick, S. M., Zonderman, A. B., and Ferrucci, L. (2011b). Hearing loss and incident dementia. *Arch. Neurol.* 68, 214–220. doi: 10.1001/archneurol.2010.362

Lin, F. R., Niparko, J. K., and Ferrucci, L. (2011c). Hearing loss prevalence in the United States. *Arch. Intern. Med.* 171, 1851–1852. doi: 10.1001/archinternmed.2011.506

Lin, F. R., Thorpe, R., Gordon-Salant, S., and Ferrucci, L. (2011d). Hearing loss prevalence and risk factors among older adults in the United States. *J. Gerontol. A Biol. Sci. Med. Sci.* 66, 582–590. doi: 10.1093/gerona/glr002

Lind, C. (2009). "Conversation repair strategies in audiologic rehabilitation," in *Adult Audiologic Rehabilitation*, eds J. Montano and J. Spitzer (San Diego, CA: Plural), 217–241.

Lind, C., Hickson, L., and Erber, N. (2010). Who said what? Sampling conversation repair behavior involving adults with acquired hearing impairment. *Semin. hear.* 31, 104–115. doi: 10.1055/s-0030-1252104

Lunner, T., Rudner, M., and Ronnberg, J. (2009). Cognition and hearing aids. *Scand J Psychol.* 50, 395–403. doi: 10.1111/j.1467-9450.2009.00742.x

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., et al. (2014). Listening effort and fatigue: what exactly are we measuring? A British society of audiology cognition in hearing special interest group 'white paper'. *Int. J. Audiol.* 53, 433–440. doi: 10.3109/14992027.2014.890296

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Mishra, S. (2014). *Exploring Cognitive Spare Capacity: Executive Processing of Degraded Speech*. Linköping: Linköping University. doi: 10.3384/diss.diva-104946

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychol. (Amst)* 27, 84–92. doi: 10.1016/0001-6918(67)90048-0

Nachtegaal, J., Festen, J. M., and Kramer, S. E. (2012). Hearing ability in working life and its relationship with sick leave and self-reported work productivity. *Ear Hear.* 33, 94–103. doi: 10.1097/AUD.0b013e318228033e

Nachtegaal, J., Smit, J. H., Smits, C., Bezemer, P. D., Van Beek, J. H., Festen, J. M., et al. (2009). The association between hearing status and psychosocial health before the age of 70 years: results from an internet-based national survey on hearing. *Ear Hear.* 30, 302–312. doi: 10.1097/AUD.0b013e31819c6e01

Norman, D. A., and Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cogn. Psychol.* 7, 44–64. doi: 10.1016/0010-0285(75)90004-3

Obleser, J., Wise, R. J., Dresner, M. A., and Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *J. Neurosci.* 27, 2283–2289. doi: 10.1523/JNEUROSCI.4663-06.2007

Park, D. C. (1999). "The basic mechanisms accounting for age-related decline in cognitive function," in *Cognitive Aging: A Primer*, eds D. Park and N. Schwarz (Philadelphia, PA: Psychology Press Taylor and Francis Group), 3–21.

Park, D. C., and Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annu. Rev. Psychol.* 60, 173–196. doi: 10.1146/annurev.psych.59.103006.093656

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *Int. J. Audiol.* 42(Suppl. 2), S26–S32. doi: 10.3109/14992020309074641

Pichora-Fuller, M. K. (2008a). Audition and Cognition: Where Lab Meets Clinic. *ASHA Lead.* 13, 14–17. doi: 10.1044/leader.FTR2.13102008.14

Pichora-Fuller, M. K. (2008b). Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. *Int. J. Audiol.* 47(Suppl. 2), S72–S82. doi: 10.1080/14992020802307404

Pichora-Fuller, M. K., Dupuis, K., Reed, M., and Lemke, U. (2013). Helping older people with cognitive decline communicate: hearing aids as part of a broader rehabilitation approach. *Semin. Hear.* 34, 308–330. doi: 10.1055/s-0033-1356643

Pichora-Fuller, M. K., and Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends Amplif.* 10, 29–59. doi: 10.1177/108471380601000103

Picou, E. M., Ricketts, T. A., and Hornsby, B. W. (2011). Visual cues and listening effort: individual variability. *J. Speech Lang. Hear. Res.* 54, 1416–1430. doi: 10.1044/1092-4388(2011/10-0154)

Posner, M. I., and DiGirolamo, G. J. (1998). "Executive attention: conflict, target detection, and cognitive control," in *The Attentive Brain*, ed. R. Parasuraman (Cambridge, MA: MIT Press), 401–423.

Pronk, M., Deeg, D. J., Smits, C., Van Tilburg, T. G., Kuik, D. J., Festen, J. M., et al. (2011). Prospective effects of hearing status on loneliness and depression in older persons: identification of subgroups. *Int. J. Audiol.* 50, 887–896. doi: 10.3109/14992027.2011.599871

Raz, N. (2005). "The aging brain observed *in vivo*: different changes and their modifiers," in *Cognitive Neuroscience of Aging. Linking cognitive and Cerebral Aging*, eds R. Cabeza, L. Nyberg, and D. Park (New York, NY: Oxford University Press), 19–57.

Ronnberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Ronnberg, J., Rudner, M., and Lunner, T. (2014). Comments from Dr. Jerker Ronnberg, Mary Rudner, Thomas Lunner. *Int. J. Audiol.* 53, 441–442; discussion 444–445.

Rudner, M., Ng, H. N. E., Rönnberg, N., Mishra, S., Rönnberg, J., Lunner, T., et al. (2011). Cognitive spare capacity as a measure of listening effort. *J. Hear. Sci.* 1, 47–49.

Rudner, M., and Lunner, T. (2014). Cognitive spare capacity and speech communication: a narrative overview. *Biomed Res. Int.* 2014, 869726. doi: 10.1155/2014/869726

Salthouse, T. A. (1992). Influence of processing speed on adult age differences in working memory. *Acta Psychol. (Amst)* 79, 155–170. doi: 10.1016/0001-6918(92)90030-H

Scarinci, N., Worrall, L., and Hickson, L. (2009). The ICF and third-party disability: its application to spouses of older people with hearing impairment. *Disabil. Rehabil.* 31, 2088–2100. doi: 10.3109/09638280902927028

Scarinci, N., Worrall, L., and Hickson, L. (2012). Factors associated with third-party disability in spouses of older people with hearing impairment. *Ear Hear.* 33, 698–708. doi: 10.1097/AUD.0b013e31825aab39

Schneider, B., Pichora-Fuller, K., and Daneman, M. (2010). "Effects of senescent changes in audition and cognition on spoken language comprehension," in *The Aging Auditory System*, eds S. Gordon-Salant, R. D. Frisina, A. N. Popper, and R. R. Fay (New York, NY: Springer), 167–210.

Shinn-Cunningham, B. G. (2007). "Why hearing impairment may degrade selective attention," in *Auditory Signal Processing in Hearing Impaired Listeners*, eds T. Dau, J. M. Buchholz, J. M. Harte, and T. U. Christiansen (Denmark: Marienlyst), 29–31.

Shinn-Cunningham, B. G., and Best, V. (2008). Selective attention in normal and impaired hearing. *Trends Amplif.* 12, 283–299. doi: 10.1177/1084713808325306

Stam, M., Kostense, P. J., Festen, J. M., and Kramer, S. E. (2013). The relationship between hearing status and the participation in different categories of work: demographics. *Work* 46, 207–219.

Stam, M., Kostense, P. J., Lemke, U., Merkus, P., Smit, J. H., Festen, J. M., et al. (2014). Comorbidity in adults with hearing difficulties: which chronic medical conditions are related to hearing impairment? *Int. J. Audiol.* 53, 392–401. doi: 10.3109/14992027.2013.879340

Stephens, D., and Jones, L. (2005). *The impact of genetic hearing impairment*. London: Whurr.

Stephens, D., and Kramer, S. E. (2009). *Living with Hearing Difficulties: The process of enablement*. Chichester: Wiley.

Strawbridge, W. J., Wallhagen, M. I., Shema, S. J., and Kaplan, G. A. (2000). Negative consequences of hearing impairment in old age: a longitudinal analysis. *Gerontologist* 40, 320–326. doi: 10.1093/geront/40.3.320

Tesser, A., and Schwarz, N. (2000). *Blackwell Handbook of Social Psychology: Intraindividual Processes*. Oxford: Wiley-Blackwell.

Tun, P. A. W., and Wingfield, A. (1999). One voice too many: adult age differences in language processing with different types of distracting sounds. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 54, 317–327. doi: 10.1093/geronb/54B.5.P317

Uchino, B. N. (2006). Social support and health: a review of physiological processes potentially underlying links to disease outcomes. *J. Behav. Med.* 29, 377–387. doi: 10.1007/s10865-006-9056-5

Uhlmann, R. F., Larson, E. B., Rees, T. S., Koepsell, T. D., and Duckert, L. G. (1989). Relationship of hearing impairment to dementia and cognitive dysfunction in older adults. *JAMA* 261, 1916–1919. doi: 10.1001/jama.1989.03420130084028

Ullsperger, M., and von Cramon, D.Y. (2006). "Funktionen frontaler Strukturen [Functions of frontal structures]," in *Neuropsychologie [Neuropsychology]*, 2nd Edn, eds H. O. Karnath and P. Thier (Heidelberg: Springer), 479–488.

Wickens, C. D. (2008). Multiple resources and mental workload. *Hum. Factors* 50, 449–455. doi: 10.1518/001872008X288394

Wingfield, A., (2014). Comment from Dr. Arthur Wingfield. *Int. J. Audiol.* 53, 442–444; discussion 444–445.

Wingfield, A. T., and Tun, P. A. (2001). Spoken language comprehension in older adults: interactions between sensory and cognitive change in normal aging. *Semin. Hear.* 22, 287–301. doi: 10.1055/s-2001-15632

World Health Organization. (2001). *International Classification of Functioning, Disability And Health, ICF*. Geneva: World Health Organization.

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., Van Beek, J. H., and Ronnberg, J. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear Hear.* 32:e16–e25. doi: 10.1097/AUD.0b013e318228036a

Zimprich, D., and Kurtz, T. (2013). Individual differences and predictors of forgetting in old age: the role of processing speed and working memory. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 20, 195–219. doi: 10.1080/13825585.2012.690364

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Lemke and Scherpiet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

REVIEW
published: 11 June 2015
doi: 10.3389/fpsyg.2015.00684

Cognitive aging and hearing acuity: modeling spoken language comprehension

Arthur Wingfield, Nicole M. Amichetti and Amanda Lash*

Volen National Center for Complex Systems, Brandeis University, Waltham, MA, USA

The comprehension of spoken language has been characterized by a number of "local" theories that have focused on specific aspects of the task: models of word recognition, models of selective attention, accounts of thematic role assignment at the sentence level, and so forth. The ease of language understanding (ELU) model (Rönnberg et al., 2013) stands as one of the few attempts to offer a fully encompassing framework for language understanding. In this paper we discuss interactions between perceptual, linguistic, and cognitive factors in spoken language understanding. Central to our presentation is an examination of aspects of the ELU model that apply especially to spoken language comprehension in adult aging, where speed of processing, working memory capacity, and hearing acuity are often compromised. We discuss, in relation to the ELU model, conceptions of working memory and its capacity limitations, the use of linguistic context to aid in speech recognition and the importance of inhibitory control, and language comprehension at the sentence level. Throughout this paper we offer a constructive look at the ELU model; where it is strong and where there are gaps to be filled.

Keywords: speech recognition, working memory, inhibition, sentence comprehension, ELU model

OPEN ACCESS

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Steve Majerus,
Université de Liège, Belgium
Jerker Rönnberg,
Linköping University, Sweden

****Correspondence:***
Arthur Wingfield,
Volen National Center for Complex Systems, Brandeis University,
415 South Street, Waltham,
MA 02454, USA
wingfield@brandeis.edu

Specialty section:
This article was submitted to Auditory Cognitive Neuroscience, a section of the journal Frontiers in Psychology

Received: *06 March 2015*
Accepted: *10 May 2015*
Published: *11 June 2015*

Citation:
Wingfield A, Amichetti NM and Lash A (2015) Cognitive aging and hearing acuity: modeling spoken language comprehension. Front. Psychol. 6:684. doi: 10.3389/fpsyg.2015.00684

Introduction

Raymond Carhart has been credited with coining the term "audiology" (an interesting mix of Latin and Greek roots), and offering the first formal course with that name at Northwestern University in 1946. In its early beginnings the issue of cognition played no role in research or teaching on hearing loss. In Newby's (1958) then-classic text in audiology, for example, the focus was squarely on peripheral hearing loss; any issues related to the pathways from the brain stem to and including the cortex was cited as the domain of neurology (Newby, 1958, pp. 53–55). Indeed, beyond supplying a definition of presbycusis as an age-related hearing loss, adult aging received no additional attention.

It is now well recognized that older adults' success in speech recognition, especially under difficult listening conditions, will be affected by cognitive factors: either in a positive way through support from linguistic context, or in a negative way where performance can be constrained by limitations in working memory and executive resources (van Rooij and Plomp, 1992; Humes, 1996; Gordon-Salant and Fitzgibbons, 1997; Wingfield and Tun, 2001; Pichora-Fuller, 2003). Just as audiology has begun to recognize that cognitive factors may play a role in performance, so cognitive psychologists engaged in research on language comprehension in older adults are beginning to recognize that the full picture of language comprehension cannot be understood without attending to the auditory declines that are common in normal aging. The joining of these two areas of expertise has seen a dramatic increase, giving rise to such terms as "cognitive hearing science" (Arlinger et al., 2009) and " cognitive audiology " (Jerger, cited in Fabry, 2011, p. 20). The introduction of these terms reflects

an increasing emphasis on the importance of taking into account how cognitive processes interact with hearing acuity in communicative behavior and remediation strategies to deal with hearing loss.

The broad sweep of issues underlying sensory-cognitive interactions in the perception and comprehension of speech raises the need for a unifying framework to guide present and near-future research. The *Ease of Language Understanding* (ELU) model (Rönnberg, 2003; Rönnberg et al., 2008, 2013) stands as such attempt. In this article we examine aspects of the ELU model that apply especially to spoken language comprehension in adult aging, where speed of processing (Salthouse, 1996), working memory capacity (Salthouse, 1994), and hearing acuity (Lethbridge-Ceijku et al., 2004) are often compromised. Throughout, we hope to offer a constructive look at the ELU model; where it is strong and where there are gaps to be filled. In so doing we use this discussion as a vehicle to examine interactions of perceptual, linguistic, and cognitive factors in spoken language understanding.

The ELU Model: A Brief Summary

The ELU model has developed from its original version (Rönnberg, 2003) to the more inclusive model as it is presented today (Rönnberg et al., 2013). The 2003 paper presents a basic framework along with a formulation to capture four parameters of spoken language understanding: (1) accuracy and features of syllable representations; (2) the speed of access to long-term memory (LTM); (3) the level of mismatch between the stimulus input and the corresponding phonology represented in the mental lexicon; and (4) the processing efficacy and storage capacity of working memory. This initial model assumed an interaction between the quality of the sensory input, information available in LTM, and the utilization of working memory. Together these would determine the ease with which language can be comprehended under difficult listening conditions. An important element in this initial presentation was a model assumption that phonological and lexical access are automatic (implicit) as long as no mismatch occurs between the sensory input and stored lexical representations. When a mismatch occurs processing becomes explicit, represented by employment of supportive context and engagement of working memory resources. This early foundation thus assumed a fundamental division between implicit and explicit components in speech understanding.

The 2013 version (Rönnberg et al., 2013) became more nuanced and more specific. In the former case it was now argued that implicit and explicit processing may operate on the interaction of phonology and semantics in parallel. As such, long-term memory (LTM) can be used either explicitly (a slow process) or implicitly (a rapid process) for understanding a spoken message. There was also an increasing attempt to say how working memory capacity relates to attention, short-term storage, inhibition, episodic LTM, and listening effort. In addition, the model in 2013 distinguishes between types of LTM (episodic and semantic) and how and when these memory systems are accessed at different stages of understanding. Rönnberg et al.'s (2008) version implied a solely feed-forward system, with the rapid and automatic multimodal binding of phonology taking place in an episodic buffer through implicit processing that matches inputs with stored representations in the mental lexicon. The 2013 version now recognizes the involvement of continuous feedback with both predictive and post-dictive (backward) feedback loops. This latter presumption is necessary given findings such as, for example, the demonstration that the perception of sub-lexical sounds are influenced by top-down word knowledge (Samuels, 2001).

Finally, in Rönnberg et al. (2013) the ELU model has been broadened to include multimodal integration in the form of visual information from seeing a talker's articulatory movements, processed in a modality-general limited capacity working memory system. In this latter regard there is certainly ample evidence for multimodal integration beginning with Sumby and Pollack's (1954) demonstration that people perceive speech in noise better when they can see the speaker's face. Access to such visual information can also be advantageous for older adults (Sommers et al., 2005; Feld and Sommers, 2009). With these recent revisions, the ELU model sets up a new line of predictions. Many of these predictions relate to the effects of different signal qualities, the type and modality of the inputs (hearing, vision, and sign language), and the relationship of working memory capacity to different encoding operations and other memory systems.

Although the ELU model has become more inclusive, there are aspects of language processing that remain underrepresented in model. We address several of these issues below. In so doing we place special emphasis on spoken language understanding by older adults following typical age-related changes in cognitive efficiency and hearing acuity. As we shall see, the cognitive literature, upon the ELU model should rely, remains unsettled on many critical issues. These issues also form a part of our discussion.

Conceptions and Control Functions in Working Memory and its Capacity

As we have noted above, working memory plays a central role in the ELU model, where it is seen as carrying a number of cognitive functions relevant to language understanding. Most conceptions of working memory in the cognitive literature have in one way or another postulated a trade-off between processing and storage, whether conceived in terms of a shared general resource (Just and Carpenter, 1992; Carpenter et al., 1994), or a limited-capacity central executive (Baddeley, 1996; Logie, 2011). Mechanisms that have been proposed to underlie the limited capacity of working memory have included time-based models in which switching attention from processing to storage or updating and refreshing the memory trace are constrained by the time parameters of these processes (Barrouillet et al., 2004, 2012). In this latter regard descriptions of working memory and executive function begin to merge, with these terms often used along with the even more general term, "resources" (often, without distinction, referred to as attentional resources, processing resources, or cognitive resources).

A model that focuses on language understanding under adverse listening conditions would benefit greatly if it could rest on settled

conceptions of working memory and executive function in the general cognitive literature. As yet such a simple consensus has yet to emerge. It might be helpful to adopt McCabe et al.'s (2010) characterization of working memory as focusing on the ability to store and manipulate information, and executive function as focusing on goal-directed behavior, monitoring and updating performance, set shifting, and inhibition (cf. Hasher and Zacks, 1988; Hasher et al., 1991; Cowan, 1999; Engle, 2002; Fisk and Sharp, 2004; Bopp and Verhaeghen, 2005; Logie, 2011), albeit with each containing elements of the other and all of these abilities associated with activity in prefrontal cortex (McCabe et al., 2010).

In its current version the ELU model cites the importance of inhibition and executive function in speech processing, but the relationship between these functions and working memory are as yet not clearly articulated within the model (Rönnberg et al., 2013, p. 10). The challenge in doing so is highlighted in McCabe et al. (2010) who report a strong correlation between tests of working memory capacity and those purported to test executive functioning ($r = 0.97$), with only processing speed showing independence. Although there is agreement that working memory capacity is limited, and more limited in older relative to younger adults (Salthouse, 1994, 1996; Salthouse et al., 2003), there is no uniform agreement within the cognitive aging literature on the mechanisms that underlie this limitation.

Our own view is closely aligned with the postulate that working memory capacity is determined by how well one can focus attention (Engle et al., 1999; Engle and Kane, 2004). A case in point is Cowan's (1999) Embedded-Process model that sees working memory as an activated subset of information within LTM. The source of the well-known capacity limitation in working memory is seen as due to the limited capacity of attentional focus that operates on the activated areas within LTM (Cowan, 1999). As such, the capacity of working memory arises from both a time limit on activation of items in memory, unless refreshed, and a limit on attentional capacity in terms of the number of items that can be concurrently activated (Cowan, 2005). What we describe here is a process-based view of working memory and working memory capacity that allows concurrent activation of representationally distributed information, a potential mechanistic account for the modality-general aspects of working memory postulated in the ELU model.

Control Functions in Working Memory

The emphasis in the ELU model is on communication, which sets it apart from many extant models of speech recognition and language understanding that focus more narrowly on specific processes and in many cases do not address how the systems operate under adverse listening conditions. Considerable research has shown that the perceptual effort attendant to poor listening conditions has a negative impact on recall of speech materials (Rabbitt, 1968, 1991; Pichora-Fuller et al., 1995; Surprenant, 1999, 2007; Wingfield et al., 2005) and comprehension of sentences that express their meaning with non-canonical word orders typical of syntactically complex speech, with this latter effect compounded by effects of age, hearing acuity, and rapid speech rates (Wingfield et al., 2006).

In the ELU model the degree of effort engendered by task difficulty affects the degree to which explicit processing will be engaged. Among such explicit processes must be an ability to monitor the ongoing capacity of working memory as speech arrives in real time. **Figure 1** shows data taken from our laboratory in which we probed the effect of listening effort on the ability to monitor the capacity of working memory as speech is arriving in real time. For this purpose we used an *interruption-and-recall* (IAR) paradigm in which participants listen to a string of recorded words with instructions to interrupt the input when they believe they have heard the maximum number of words that will allow for perfect recall of what has been heard. Germaine to our present interests, the word-lists were presented at one of two sound levels: at 25 dB SL to represent listening ease, and 10 dB SL to represent effortful listening. The participants were young adults with age-normal hearing (Amichetti et al., 2013, Experiment 2).

Figure 1A shows the mean number of words correctly recalled in a simple baseline span task in which listeners heard lists varying in length from one to 12 items for immediate recall. It can be seen that for list lengths of up to three words recall is at ceiling, and at near-ceiling for a four-item list length at both intensity levels, thus confirming the audibility of the stimuli at the two sound levels. Beyond a four-item list, additional stimulus items yield progressively smaller recall gains that never peak beyond means of 5.8 items for the 25-dB SL lists and 4.3 items for the 10-dB SL lists. This small but significant difference affirms the above-cited negative effect of effortful listening on recall.

Figures 1B,C are of greater interest as they show what happened when participants heard supra-span lists with instructions to interrupt the word-lists with a keypress when they believed they had heard the maximum number of words that they could recall with perfect accuracy. The middle panel shows the distribution of segment sizes participants selected for recall for the 25-dB SL and 10-dB SL presentation levels in this IAR condition. One can see a shift in the peaks of the two distributions, from a modal self-selected segment length of six words for lists at the louder 25-dB SL level, to seven words, at the 10-dB SL level. Specifically, at 25 dB SL the modal segment size of six words was close to the mean for accurate item recall of 5.8 words in the baseline span condition at that sound level shown in the left panel, suggesting a good ability to calibrate segment size selections with actual memory span. By contrast, in the effortful listening condition, listeners appeared to lose this close calibration. That is, a reduced memory span for accurate recall of 4.3 words for 10 dB SL lists in the baseline condition was not accompanied by listeners adaptively taking shorter segment sizes for recall in the IAR condition.

The right panel shows the number of words recalled in the IAR conditions for list lengths that had more than 10 examples on which to base a mean. The dual-task nature of the IAR condition (the listener must make continuous capacity judgments while holding what has been heard to that point in memory) reflects a greater cognitive load than in baseline span task. As would be expected if listening effort draws on already strained resources in the IAR task, while for the 25-dB SL presentations the IAR-produced spans are similar to baseline spans at 25 dB SL, recall accuracy for the IAR spans at the more effortful 10 dB SL level were

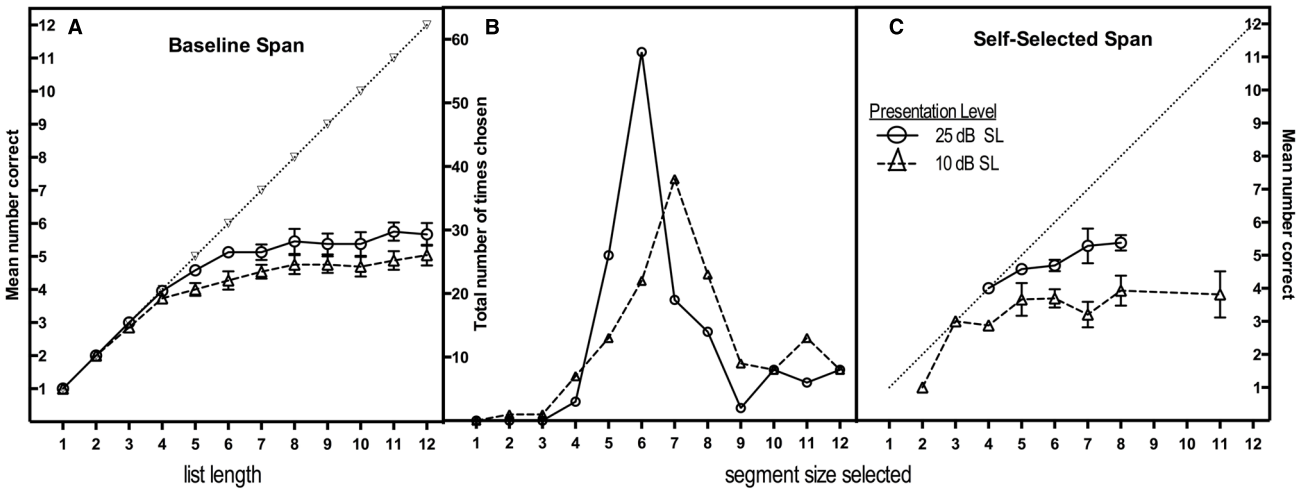


FIGURE 1 | Mean number of words correctly recalled as a function of the number of words presented (baseline span) for words presented at either 25 dB SL or 10 dB SL (A); the distribution of segment sizes selected in the interruption-and-recall (IAR) condition at the two sound levels (B); and the number of words recalled as a function of the number of words selected in the IAR condition at the two presentation levels (C). Error bars in left and right panels represent one standard error. (From **Figure 2** in Amichetti et al., 2013, Copyright 2013 by the Psychonomic Society. Reprinted with permission.)

reduced relative to the corresponding baseline span presented at 10 dB SL.

As we have noted, the ELU model asserts that a degraded (perceptually effortful) signal leads to a shift from automatic to controlled processing with an engagement of working memory resources. We show with the above data that this control itself may be affected by the necessity to process a low-quality signal. In part the lower sound level may have slowed the stimulus encoding, resulting in an overlap in time in which the cognitive system is concurrently conducting perceptual and encoding operations on one stimulus as another is arriving (Miller and Wingfield, 2010; Piquado et al., 2010). It is also possible that a reduced stimulus intensity may truncate the duration of an already rapidly fading echoic trace (Baldwin, 2007; Baldwin and Ash, 2011).

This control function in working memory may be obscured in natural speech if listeners are allowed to periodically interrupt a spoken narrative to give themselves time to process what they have heard before the arrival of yet more information. In this case both young and older adults tend to interrupt the speech input at major linguistic clauses and sentence boundaries rather than after a set number of words (Wingfield and Lindfield, 1995; Piquado et al., 2012; see also Wingfield et al., 1999; Fallon et al., 2006). Importantly, such findings are indicative of listeners' access to syntactic and semantic knowledge as the speech is being heard, and hence being involved in very early stage processing. We will address the implications of early access in several places in the following discussion.

The Implicit versus Explicit Distinction

Fundamental to the ELU model is the position that when speech quality is good, with a clear match between acoustic input and its corresponding phonological representation in LTM, lexical recognition will be automatic ("implicit"). That is, lexical access will be rapid, resource-free, and will not require access to top-down information such as linguistic or semantic context. When the input quality is poor, whether due to external factors such as background noise, or internal factors such as hearing loss or a distorted phonological representation in LTM consequent to a long-term hearing impairment, the degraded information can be supplemented by linguistic or real-world knowledge, a process that requires explicit or "effortful conscious processing" (Mishra et al., 2013, p. 2).

Use of the terms *implicit* and *explicit* processing in the ELU model resonate with the early (LaBerge and Samuels, 1974; Shiffrin and Schneider, 1977), but still often used distinction in the cognitive literature between *automatic* versus *controlled processes*. In the context of speech perception, automatic processes emphasize bottom-up, stimulus-driven processing that is rapid, obligatory, and demanding few if any resources. By contrast, controlled processes tend to be top-down, voluntary, and to one degree or another resource-demanding (Pashler et al., 2001). They are also assumed to require some level of awareness (LaBerge and Samuels, 1974; Posner and Snyder, 1975; Shiffrin and Schneider, 1977; Flores d'Arcais, 1987). All of these attributes fit squarely with the characterization of implicit and explicit processing as represented in the ELU model.

Although early-stage perception is often considered to be automatic, arguments have been offered for cognitive and attentional control operating at the earliest stages of input processing of speech (Nusbaum and Magnuson, 1997; Heald and Nusbaum, 2014). It should also be recognized that a system that appears to be resource-free could require resources but not those shared with other processes. This exact position was taken by Caplan and Waters (1999) who argued that on-line syntactic operations are conducted by sentence-specific resources not measured by traditional working memory tasks such as the Daneman and Carpenter (1980) readings span task or its several

variants. They suggest that the appearance of effects working memory limitations on sentence processing represent post-interpretive processes rather than on initial syntactic parsing. Our present focus, however, is the specific assertion in the ELU model that when there is degraded input perceptual operations will shift from automatic to controlled processing, with the latter increasing the drain on working memory resources (Rönnberg et al., 2013).

Proposals of binary, either-or process distinctions have been a hallmark of early theory development in cognitive psychology such as distinctions drawn between semantic versus episodic memory (Tulving, 1972), procedural versus non-procedural learning (Squire, 1994), implicit versus explicit memory in reference to priming studies (Schacter, 1987), and so forth. In each case subsequent studies have shown none of these proposed distinctions to be process pure. In a similar way, the distinction between automatic (implicit) versus controlled (explicit) processes can best be seen as two ends of a continuum and a matter of degree rather than the sharp contrast current in the ELU model.

Although drawing a distinction between implicit and explicit processes, Rönnberg et al. (2013) note that the extent to which explicit or implicit processing may be employed can vary over the course of a single task, with the ratio changing from moment to moment during a conversation depending on signal quality and speech content (see also Rönnberg et al., 2010).

It is the case that the automatic versus controlled distinction retains descriptive utility (Birnboim, 2003; Schneider and Chen, 2003), but only insofar as one thinks of some operations being potentially "more automatic" than others in a relative or graded sense (Chun et al., 2011).

The Match versus Mismatch Distinction

The match versus mismatch distinction highlighted in the ELU model may be accepted as an idealized principle, although such a distinction should be treated with caution. This is so because there is rarely a perfect match between a phonological input and the phonological representation of an item in the mental lexicon. This is due to the variability in the way words and their sub-lexical elements are articulated from speaker to speaker, and effects of syllabic context within a single speaker (Liberman et al., 1967; Mullennix et al., 1989).

At the more cognitive level, analyses of natural speech show that speakers tend spontaneously to employ a *functional adaptation* in their production. That is, we tend to articulate more clearly words that cannot be easily inferred from context, and to articulate less clearly those that can (Hunnicutt, 1985; Lindblom et al., 1992). It is not assumed that these dynamic adjustments are consciously applied by the speaker, any more than we assume that listeners are necessarily consciously aware of using acoustic and linguistic context in their perceptual operations.

Because of this functional adaptation, what one might call an articulatory *principle of least effort*, words are often under-articulated when they can be predicted from the context, and many words would be unintelligible were it not for the phonemic and linguistic context in which they are ordinarily heard (Lieberman, 1963; Pollack and Pickett, 1963; Grosjean, 1985; Wingfield et al., 1994). Because of this variability perfect-match template matching models of perception must be an ineffective account of perceptual identification. To the extent that the ELU model presumes a perfect or near perfect match between phonological inputs and stored counterparts in LTM as the default condition with natural speech, this would be out of tune with these data. It should be noted that although the early Rönnberg (2003) formulation implied a stark contrast between a perfect match versus one that requires top-down support, the current model version sees word recognition in terms of a threshold function affected by phonological and semantic attributes (Rönnberg et al., 2013). This question relates to broader issues in the role of linguistic context in speech recognition and comprehension.

The Role of Context

A common view in speech recognition is that questions related to effects of context should be framed in terms of top-down effects operating on initially stimulus-driven perceptual processes. The ELU model is in general accord with this principle, although an apparently conflicting observation appears in the suggestion in Rönnberg et al. (2013) that if a sentence context is sufficiently predictive, a target word might be activated even with minimal phonological input (Rönnberg et al., 2013). This presumption, although consistent with everyday experience, would not seem to follow at first look from the precepts of the ELU model. It would follow, however, from a number of extant models of word recognition.

Most models of word recognition, to include the ELU model, assume a reciprocal balance between bottom-up information determined by the clarity of the speech signal and top-down information supplied by a system of linguistic knowledge (e.g., Morton, 1969, 1979; McClelland and Elman, 1986; Marslen-Wilson, 1987). It is the compensatory availability of preserved linguistic knowledge and the procedural rules for its use that accounts for the general effectiveness of speech comprehension in adult aging in spite of cognitive and sensory declines (Wingfield et al., 1991; Pichora-Fuller et al., 1995; Wingfield and Stine-Morrow, 2000; Pichora-Fuller, 2003). Although these principles are embodied within the broad outlines of the ELU model, questions remain as to whether context comes into play before, during, or after the acoustic representation of a word unfolds in time.

A model that assumes that context activates lexical possibilities before a stimulus word is heard was embodied in one of the earliest interactive models: the so-called "logogen" model that also went through a period of development (Morton, 1964a,b, 1969, 1979). Morton postulated a "dictionary" of "units" (later re-named "logogens"), with each unit corresponding to a word represented in LTM. When the level of activation of a logogen exceeds a critical level, the unit "fires," and the corresponding word is available as a response.

In this model each unit has a resting potential, or base level of activation, determined by the relative frequency with which the unit has fired in the past. This is reflected behaviorally in the *word frequency effect*, in which words that have a high frequency of occurrence in the language are recognized faster or with less

stimulus information than low-frequency words (Howes, 1957; Grosjean, 1980). Following the firing of a unit its resting level of activation increases sharply, resulting in recency or repetition priming, and then decays slowly. Through direct connections with other units, the activation of any given unit adds to the level of activation of all associated units, whether this association is semantic, categorical, or based on shared attributes.

In operation, a sensory input would be coded in terms of the presence of detected phonological features, the presence of which would simultaneously increase the level of activation of all units sharing these phonological features. Thus, the unit sharing the greatest number of features with the presented stimulus would receive the greatest increase in its level of activation. It can be seen from this formulation that the amount of stimulus information required for a unit to exceed its critical level and "fire," would be lower either when there is already a high level of residual activation (the word frequency effect), when the level of activation has been temporarily raised by a recent firing of the unit (recency priming), or by the firing of an associated unit or units (an effect of context).

Within the Logogen model, a highly constraining linguistic or environmental context that increases the likelihood of occurrence of a stimulus word will increase the level of activation of that item in the mental lexicon, thus priming the entry even before the stimulus is actually encountered. The higher the level of activation, the less stimulus information will be required for recognition of the target word. Activation due to contextual expectancy would thus override units' initial resting potentials initially determined by their relative frequency of occurrence in the language, and hence, their likelihood of re-occurrence. A constraining linguistic or environmental context would also override other factors known to affect the intelligibility of individual words, such as the detrimental effect of a large number of words that share initial or overall phonology with the target word (cf. Tyler, 1984; Wayland et al., 1989; Wingfield et al., 1997; Luce and Pisoni, 1998). These general principles have been embodied in a number of models, to include TRACE, a computational model in which the above factors, operating in parallel, can be implemented by transient weighting factors (McClelland and Elman, 1986).

A correlate of Morton's model is that if the level of activation of a lexical unit is sufficiently raised due to a high probability of it being encountered, a lexical unit may "fire" in the absence of objective stimulus information. It can be seen that Morton's logogen model and others like it offer a mechanistic account noted by Rönnberg et al. (2013) that if a sentence context is sufficiently predictive, a target word might be activated even with minimal phonological input. This principle of an inverse relationship between the *a priori* probability of a word and the amount of phonological information needed for its recognition is a well established finding in the literature for both spoken and written words and for both young and older adults (Black, 1952; Bruce, 1958; Morton, 1964a,b; Cohen and Faulkner, 1983; Madden, 1988). It should be pointed out, of course, that the more likely scenario following the same principle is the misidentification of an indistinct word as a word with a similar sound that is a closer fit to a semantic context (Rogers et al., 2012). Either case, however, would necessitate a closer look within the ELU model at whether context raises lexical activation before (Morton, 1969), during (Marslen-Wilson and Zwitserlood, 1989), or after (Swinney, 1979) the word unfolds in time.

In contrast with models that assume that linguistic context raises target activation even prior to acoustic input, we have seen that a basic tenet of the ELU model is that an acoustically clear stimulus with a correspondingly rich mental representation results in automatic (implicit) lexical access; a rapid, obligatory, resource-free process. In the model context comes into play only when poor stimulus quality does not allow an immediate match at which point context "kicks in." The process being described is suggestive of early modular models of lexical access such as Forster's (1976; 1981) argument for autonomous lexical access: a self-contained modular system, with restricted access to information. Such an "informationally encapsulated" (context-free) process fit within Fodor's (1983) broader argument for modularity within cognitive domains and processes.

The positive influence of a constraining sentence context or other sources of semantic priming on the accuracy or speed of lexical access (e.g., Holcomb and Neville, 1990) appears as inconsistent with the postulate of a context-impenetrable modular view of lexical access. This issue is not easily settled in spite of a history of creative experiments intended to determine whether the facilitation observed with a constraining sentence context reflect a true access effect (cf., Swinney, 1979; Seidenberg et al., 1982, 1984; Stanovich and West, 1983).

The issue is whether the well-documented effects of expectancy on ease of lexical access, and especially the suggestion that a sufficiently strong expectation can activate a lexical entry in the absence of sensory input, is most compatible with a pre-lexical (e.g., Morton, 1969) or a post-lexical (e.g., Forster, 1981) effect. Our reading of the ELU model appears to favor both positions, an issue that would need to be reconciled as the model develops in detail.

Before leaving this issue, we might also suggest that a complete model for word recognition should include not only the level of activation of a lexical entry as determined by contextual expectancy and the goodness of fit with the stimulus, but also on the individual's acceptance criterion level. This flexible criterion level would be determined by such factors as the priority given to speed versus accuracy (Wagenmakers et al., 2008) or the reward for a correct recognition versus the negative consequences of making an erroneous identification (Green and Swets, 1966). This position thus adds motivational state to the quality of the sensory input and the sensory capacities of the listener.

Age and Inhibition in Word Recognition: The Role of Working Memory

Benichov et al. (2012) examined ease of recognition of sentence-final words heard in noise with participants aged 19–89 years, with levels of hearing acuity ranging from normal hearing to mild-to-moderate hearing loss. Regression analyses showed that hearing acuity, although a predictor of the signal to noise ratio necessary to correctly recognize a word in the absence of a constraining linguistic context, dropped away as a significant

contributor to recognition of sentence-final words by the time the linguistic context was strongly predictive. By contrast, a cognitive composite of individuals' episodic memory, working memory, and processing speed accounted for a significant amount of the variance in word recognition for words heard in a neutral context and for all degrees of contextual constraint examined. (The contextual probability of the target words was taken from published "cloze" norms, which report the percentage of participants who give particular words when asked to complete sentence stems with the final word missing.)

One likely candidate for the role that working memory capacity may play in word recognition was revealed in a study by Lash et al. (2013) who examined effects of age, hearing acuity, and expectations for the occurrence of a word based on a linguistic context. Importantly, the study also examined the effects of competition from other words that might also fit the semantic contexts. Lash et al. (2013) used the technique of *word-onset gating*, in which a listener is presented with an increasing amount of a word's onset duration until the word can be correctly identified (Grosjean, 1980, 1996). When a linguistic context is absent, word recognition is affected by the number of words that share the initial sounds with the target word (Tyler, 1984; Wayland et al., 1989), further limited by words that share syllabic stress (Wingfield et al., 1997; Lindfield et al., 1999; see also Wingfield et al., 1990).

A major focus of the Lash et al. (2013) study was the effect of a linguistic context on word recognition that, as we have previously indicated, will override such factors as word frequency or the number ("density") of phonological competitors as determinants of word recognition. A critical feature of published cloze norms (e.g., Lahar et al., 2004), however, is that when participants have been asked to complete sentence stems, also reported is the full range of responses given by each of the participants, and the number of participants giving these alternative responses. These data allow one to estimate not only the expectancy of a sentence-final word based on the transitional probability of that word in the sentence context, but also the uncertainty (*entropy*) implied by the number, and probability distribution, of alternative responses that also might be implied by the context. Lash et al. (2013) found that while both young and older adults' word recognition benefitted from a sentence context that increased word expectancy, a differentially negative effect of the presence of strong competitor responses was found for older adults independent of hearing acuity.

This latter finding is consistent with Sommers and Danielson's (1999) proposition that older adults have greater difficulty than their young adult counterparts in inhibiting non-target responses. In Sommers and Danielson's (1999) case the competition came from the presence of a larger number of phonological "neighbors" of target words. The present case differed only in that response competition came from the distribution of words that also shared a contextual fit with a semantic context. Such results would be expected from arguments that older adults have a general inhibition deficit (Hasher and Zacks, 1988), that in this case, would interfere with word recognition.

A subsequent study by Lash and Wingfield (2014) directly examined working memory capacity and effectiveness of inhibition in word recognition as would be predicted from observations present in the current version of the ELU model. This study was based on the finding that gradually increasing the clarity of a stimulus until it can be correctly identified retards its recognition relative to when a stimulus is presented just once, even at a level of clarity below that needed for recognition using an ascending presentation. This finding, observed originally for degraded visual stimuli, has been interpreted as reflecting the negative effect of interference from incorrect identification hypotheses formed during the incremental presentations that would not be present with a single presentation (Bruner and Potter, 1964; Snodgrass and Hirshman, 1991; Luo and Snodgrass, 1994).

Lash and Wingfield (2014) conducted an analogous study for spoken words using word-onset gating with older adults ($M = 75$ years) with good hearing acuity (PTA < 25 dB HL) and an age-matched group with a mild-to-moderate hearing loss. A group of young adults with normal hearing acuity was also included for comparison. For each individual we determined the word-onset gate size that allowed the participant to recognize correctly 40 to 60% of target words when they were presented successively with increasing onset durations (an *ascending presentation*). We also determined for each individual the recognition accuracy level for comparable words presented just once (a *fixed presentation*) with the same gate size that yielded the 40 to 60% correct recognition with an ascending presentation. The size of the interference effect from ascending presentations would be indexed by the difference between word identification rates under the two presentation conditions. The question was whether individual differences in working memory capacity might predict one's ability to inhibit interference from false identification hypotheses presumed to be formed in the course of the incrementally larger and larger word onset durations represented in the ascending presentation condition (e.g., Snodgrass and Hirshman, 1991; Luo and Snodgrass, 1994).

As might be expected from age and inhibition arguments, the older adults in the study showed a larger interference effect from ascending presentations than the young adults. Germaine to our present question, a follow-up regression analysis revealed that participants' reading spans, taken as a measure of working memory capacity (Daneman and Carpenter, 1980; McCabe et al., 2010), contributed significantly to the size of the interference effect (see Lash and Wingfield, 2014, for full details). The reading span test, which we discuss in a subsequent section, was used rather than a listening span version (e.g., Wingfield et al., 1988) to avoid a potential confound with hearing acuity.

This effect of working memory span on the effectiveness of inhibition can be illustrated most clearly in **Figure 2** in which we have taken data from Lash and Wingfield (2014) and have plotted the percentage of correct identifications for the same gate size when words were presented in the fixed versus the ascending presentation conditions separated by participants' working memory span. A participant was considered to have a high working memory span (left panel) if they scored greater than one standard deviation above the mean for their age cohort determined by McCabe et al. (2010), or a low span if they did not (right panel). These data are based on a subset of participants from Lash and Wingfield (2014) where high and low span participants within each participant group were equal in number and matched for age.

Although for the high span participants some variability appears in the difference between identification scores for the fixed versus ascending presentation conditions, especially for the young adults, none of these differences reached significance. By contrast, the lower span participants in each of the three participant groups consistently show a significant interference effect even after adjusting for differences in baseline recognition accuracy.

These data can thus be taken to offer empirical support for the suggestion in Rönnberg et al. (2013) that working memory capacity may affect the efficiency of inhibitory processes (see also Sörqvist et al., 2012). It should be noted, however, that a relationship between working memory capacity and effectiveness of inhibition leaves open the direction of causality. Indeed, an influential argument has been made that it is a failure of the ability to inhibit off-target interference that may determine the size of one's working memory capacity (Hasher and Zacks, 1988; Hasher et al., 1991). We will have more to say on this topic in the following section.

Input Challenge at the Sentence Level: Deep versus Shallow Processing

A premise of the ELU model is that a perceptual mismatch due to a poor quality stimulus causes a shift from implicit (automatic) to explicit (controlled) processing where support from linguistic or environmental context are brought into play through involvement of working memory. As outlined in the model, this shift will slow processing but hopefully lead to a successful solution. Because syntactic resolution of a sentence is arguably a precursor to determination of sentence meaning, this would imply that, when speech quality is poor, listeners will engage in an especially detailed and explicit syntactic analysis. Rönnberg et al. (2013), however, offer a qualification: when placed under time pressure, and if the listener is willing to accept the gist of the message, such a close analysis might not take place (Rönnberg et al., 2013, p. 10).

There is no doubt that this latter point is true, both intuitively and empirically. We would suggest, however, that in natural language comprehension such gist processing may be the rule rather than the exception. This would be so since in listening to spoken discourse one is almost always under time pressure due to the rapidity of natural speech and the transient nature of the speech signal. Ordinary speech rates average between 140 to 180 words per minute, and can often reach 210 words per minute as, for example, with a radio of TV newsreader working from a prepared script (Stine et al., 1990).

Although in many cases a complete syntactic analysis may be conducted as a precursor to determining a sentence meaning, there is considerable evidence that listeners often, perhaps more often, take processing short-cuts, sampling key words and using plausibility to understand the meaning of an utterance. Because we live in a plausible world this strategy will in most cases yield rapid and successful comprehension, albeit with comprehension errors should one encounter a sentence with an unexpected or implausible meaning.

Analyses of everyday discourse show that most of our sentences, when they are in fact grammatical, tend to have meaning expressed in a relatively simple noun-verb-noun canonical word order with the first word representing the agent or source of the action (Goldman-Eisler, 1968). Thus, so long as the syntax is represented by canonical word order and the meaning of a sentence is plausible, a gist analysis will most often yield a correct understanding. This strategy goes unnoticed because it invariably works; it is revealed, however, when comprehension fails. In such cases listeners "mishear" a sentence as if it were sensible, such as the sentence, "The teenager that the miniskirt wore horrified the mother" (Stromswold et al., 1996). Examination of individuals' comprehension of such sentences have shown that comprehension errors frequently occur, suggesting the absence of a full syntactic analysis of a sentence input in favor of sampling key words, assuming that the word order represents the meaning in a canonical form, and that the semantic relations being expressed in the sentence are plausible (Fillenbaum, 1974; Sanford and Sturt, 2002).

Ferreira (2003) has formalized these notions, suggesting that heuristic short-cuts may be taken by all listeners, by-passing a full syntactic analysis but instead using word-order and plausibility as a rapid first-pass comprehension strategy (Ferreira et al., 2002; Ferreira, 2003; Ferreira and Patson, 2007). As Ferreira et al. (2002) have argued, it should not be assumed that all relevant information from a detailed and time-consuming lexical and syntactic analysis will be used in everyday comprehension. Sanford and Sturt (2002), from the perspective of computational linguistics, come to a similar conclusion. That is, to use Ferreira and Patson's (2007) words, sentence processing is as often as not conducted at a level of analysis that is "good enough" for comprehension. As we have argued above, this processing strategy will yield the right answer more often than not. It is consistent with the slowed processing and limited working memory capacity of older adults that Christianson et al. (2006) have argued that a "good enough" processing heuristic may be even more common in the elderly.

Working Memory and Language Comprehension

There are a variety of working memory measures in the literature designed to capture operational capacity. Important among them is the reading span task introduced by Daneman and Carpenter (1980), that focuses more specifically on verbal working memory (Carpenter et al., 1994). It is a version of this reading span task that serves as the preferred measure of working memory in the ELU-related studies conducted by the Rönnberg group.

The reading span (or listening span) task requires the listener to read (or listen to) a series of sentences and, to insure the sentences are being comprehended, to state after each sentence whether it is true or false, or in some variants, whether the meaning of the sentence is plausible or implausible. After a set of sentences is finished the reader (or listener) must recall the final word of each sentence, or he or she receives a signal to recall either the last word or the first word of each of the sentences. The span is taken as the number of sentences that allow accurate recall of the final, or the first or final words depending on the version (cf. Daneman and Carpenter, 1980; Rönnberg et al., 1989; Waters and Caplan, 1996; McCabe et al., 2010). As previously noted, the reading span,

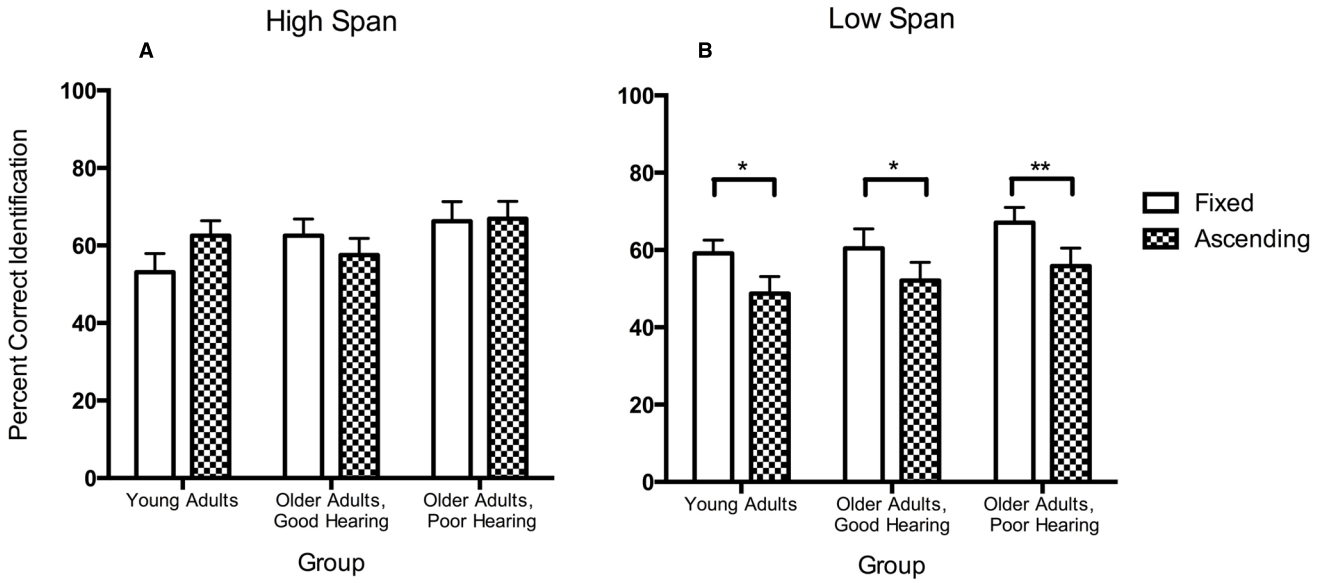


FIGURE 2 | Percentage of words correctly identified with the same onset gate size when stimuli were presented under fixed versus ascending procedures for young adults with age-normal hearing acuity and older adults with good hearing acuity or a mild-to-moderate hearing loss.

Panel **(A)** shows participants with high working memory spans. Panel **(B)** shows these data for participants with lower working memory spans. Error bars represent one standard error. (Data from Lash and Wingfield, 2014, Psychology and Aging, Viol. 29.) $^{*}p < 0.05$, $^{**}p < 0.01$.

as opposed to a listening span version, is preferable when speech is involved in order to avoid a confound with hearing acuity or stimulus clarity.

We earlier cited the claim by Caplan and Waters (1999), based on their work and the work of others, that working memory, at least as tested with the reading span task of Daneman and Carpenter (1980) and its variants, does not constrain, or by inference carry, on-line sentence comprehension. In contrast, the well-known meta-analysis by Daneman and Merikle (1996) showed reading span scores to reliably predict performance on a number of language comprehension and language memory tasks.

In addition to mixed findings in experimental studies relating reading spans to efficacy in language comprehension (see, the review in Wingfield et al., 1998) there is a similar case for the ability of working memory span as measured by reading span, as a predictor of perception of speech in noise or with reduced hearing acuity (cf. Akeroyd, 2008; Schoof and Rosen, 2014; Füllgrabe et al., 2015).

It is possible that the mixed findings in studies using the reading span as a measure of verbal working memory may lie in the intentional complexity of the reading span task itself, with this complexity allowing task demands or nuances of the instructions to affect the sensitivity of the span scores across different experiments. When one considers the reading span task it can be seen that there is an opportunity for a trade-off on the part of the reader or listener between recalling the sentence-final or sentence-initial words versus processing efficiency on the sentence comprehension component of the task. Indeed, individual differences in strategy use and session-to-session variability has been shown to occur in even less complex memory tasks (e.g., Logie et al., 1996).

Waters and Caplan (1996) recognized that the reading span task, because it involves both storage and processing components, is a better measure of working memory than a simple span test that has only a storage component. The task also has face validity as both the reading span task and language comprehension require temporary storage of verbal material along with ongoing syntactic and semantic computations. As Waters and Caplan note, this complexity of the Daneman and Carpenter (1980) reading span task focuses solely on the storage component of the task (recalling the sentence final words as the span measure) but not the efficiency with which the sentence comprehension component is conducted. To overcome this limitation they suggest a more valid measure might be represented by an index that takes into account sentence comprehension accuracy, the number of sentence final words that can be recalled, and as a measure of efficiency at sentence processing, response times to the sentence judgments. Represented as a z-score they show this composite measure to have better test-retest reliability than the original Daneman and Carpenter (1980) span test.

An additional criticism of the Daneman and Carpenter (1980) span test is that participants always know in advance that they will be asked to recall the last word of each of the sentences. That knowledge might lead to development of processing strategies by the participant. To overcome this issue Rönnberg et al. (1989) developed a span task that uses a post-cueing method in which the participant reads the stimulus sentences without knowing in advance whether they will be asked to recall the first or the final word of each sentence. This instruction is given after a sentence set has been presented.

In these regards, we suggest that a large-scale meta-analysis of studies compare and contrast findings using extant variations of the reading span task. Such an analysis should include relative strengths in terms of test-retest reliability where available.

The above discussion has focused more on the reading span as a measure of working memory capacity than on the memory systems that may be involved in speech comprehension at the sentence level. On the one hand, our discussion of "good enough" sentence processing suggests that an abstract representation of sentence meaning is formed as a sentence is being heard. On the other hand, our ability to "replay" the sensory input to retroactively repair an initial misanalysis of a garden-path sentence implies the support of a briefly sustained veridical trace of the input.

This apparent paradox was recognized by Potter (1993; Potter and Lombardi, 1990), who proposed that as a sentence is heard, both a verbatim trace of the spoken input and a semantic abstraction are concurrently formed and briefly stored in memory. Depending on the momentary needs of the listener or complexity of the speech materials, the individual might rely more or less heavily on the transient verbatim trace, whether this is thought of as a phonological, articulatory, or echoic store. In everyday listening the default mode may be reliance on the abstracted semantic trace for constructing narrative coherence, with the concurrently available verbatim trace accessible for a brief period if needed for specific task requirements or if access to the original input is needed in order to rescue an initial processing error. In the case of understanding meaningful speech, such a model might account at least in part for many of the paradoxes outlined above.

Resource-Limited versus Data-Limited Processes

In performing a complex cognitive task one would expect that, at least to some limit, the level of performance will improve with the amount of effort (resources) given to that task. This refers to a task that is "resource-limited": the upper limits on performance will be set only by the amount of resources one is willing, or able, to apply to it (Norman and Bobrow, 1975). In cases of degraded input, performance can often be improved with additional effort. There are other cases where the stimuli are of such poor quality that no amount of effort or allocation of resources will improve the level of performance. In such cases, when the upper limit on performance is determined by the limited quality of the stimulus, the task can be referred to as data-limited (Norman and Bobrow, 1975). Most tasks, even ones with a poor quality stimulus, are resource-limited up to some point where one's performance is limited only by the amount of resources one is willing to devote to it. It is only beyond this point that one can say that the task is data-limited. Although questions have arisen about distinguishing between a data-limited transition and possible constraints of a ceiling effect (Norman and Bobrow, 1975; Kantowitz and Knight, 1976). Norman and Brobrow's(1975) conceptualization is a descriptively important one.

Within the context of what Norman and Bobrow (1975) would call the resource-limited range, one can describe three "zones" of listening conditions: (1) effortless listening, where working memory resources are not drained by perceptual processing demands, (2) effortful but successful listening where errors will occur unless resources can be reallocated from other tasks, and (3) effortful but error-prone listening which is not yet data-limited, but where there are insufficient or non-optimally allocated resources (see Schneider and Pichora-Fuller, 2000; Pichora-Fuller, 2003, for discussions). Poor-hearing older adults would reach these points of effortful listening with higher sound levels than those with better hearing, and they would be reached sooner for more complex speech materials than simpler materials.

Although traditionally theorists have focused on just one direction of activity, whether on limited resources constraining perceptual effectiveness (Kahneman, 1973; de Fockert et al., 2001; Lavie et al., 2004) or perceptual effort reducing higher-level cognitive effectiveness (Rabbitt, 1968, 1991; Dickinson and Rabbitt, 1991; Murphy et al., 2000) one can postulate a single interactive dynamic which may operate in both directions: limited resources may impede successful perception when the quality of the sensory information requires perceptual effort for success, while successful perception in the context of a degraded stimulus or a hearing loss may draw on resources that might otherwise be available for downstream cognitive operations. These notions fit acceptably within the ELU model and it is hoped that they are more fully developed in future versions of the model.

Conclusion

The ELU model can fairly be represented as a work in progress with many gaps to be filled. The model nevertheless serves as a useful framework for thinking critically about language understanding, especially under difficult listening conditions. That is, a model has value not only when it answers all of our questions, accounts for extant data, and makes specific predictions for experiments yet to be conducted. A model also has value when close scrutiny highlights what we know and what we do not know; the broader the sweep of the model the more this is likely to be so.

Our goal in this discussion has been to point to places in the model where there are gaps that are yet to be filled and where the model could be productively expanded. In doing so we acknowledge that the ELU model represents a unique attempt to formulate a unifying framework to describe sensory-cognitive interactions especially under difficult listening conditions.

An important feature in the development of the ELU model has been a shared focus both on theory and on the practical implications of cognitive resources in remediation in the case of hearing loss (e.g., Rudner and Lunner, 2013). The effectiveness of the rapid development of sophisticated signal processing algorithms, whether in traditional hearing aids or in cochlear implants, must take into account the cognitive supports and cognitive constraints of the user, especially, we suggest, in the case of the older listener. The integrative approach of the ELU model offers an ideally suited framework on which to carry continued research on this critical interaction.

Acknowledgments

Our work is supported by NIH grant R01 AG019714 from the National Institute on Aging (AW). We also acknowledge support from training grant T32 AG000204 (NMA and AL) and support from the W.M. Keck Foundation.

References

Akeroyd, M. A. (2008). Are individual differences in speech perception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Amichetti, N. M., Stanley, R. S., White, A. G., and Wingfield, A. (2013). Monitoring the capacity of working memory: executive control and effects of listening effort. *Mem. Cogn.* 41, 839–849. doi: 10.3758/s13421-013-0302-0

Arlinger, S., Lunner, T., Lyxell, B., and Pichora-Fuller, M. K. (2009). The emergence of cognitive hearing science. *Scand. J. Psychol.* 50, 371–384. doi: 10.1111/j.1467-9450.2009.00753.x

Baddeley, A. D. (1996). "The concept of working memory," in *Models of Short-Term Memory*, ed. S. Gathercole (Hove: Psychology Press), 1–28.

Baldwin, C. L. (2007). Cognitive implications of facilitating echoic persistence. *Mem. Cogn.* 35, 774–780. doi: 10.3758/BF03193314

Baldwin, C. L., and Ash, I. K. (2011). Impact of sensory acuity on auditory working memory span in young and older adults. *Psychol. Aging* 26, 85–91. doi: 10.1037/a0020360

Barrouillet, P., Bernardin, S., and Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *J. Exp. Psychol. Gen.* 133, 83–100. doi: 10.1037/0096-3445.133.1.83

Barrouillet, P., De Paepe, A., and Langerock, N. (2012). Time causes forgetting from working memory. *Psychon. Bull. Rev.* 19, 87–92. doi: 10.3758/s13423-011-0192-8

Benichov, J., Cox, L. C., Tun, P. A., and Wingfield, A. (2012). Word recognition within a linguistic context: effects of age, hearing acuity, verbal ability and cognitive function. *Ear Hear.* 33, 250–256. doi: 10.1097/AUD.0b013e31822f680f

Birnboim, S. (2003). The automatic and controlled information-processing dissociation: is it still relevant? *Neuropsychol. Rev.* 13, 19–31. doi: 10.1023/A:1022348506064

Black, J. W. (1952). Accompaniments of word intelligibility. *J. Speech Hear. Disord.* 17, 409–418. doi: 10.1044/jshd.1704.409

Bopp, K. L., and Verhaeghen, P. (2005). Aging and verbal memory span: a meta-analysis. *J. Gerontol. Psychol. Sci.* 60B, P223–P233. doi: 10.1093/geronb/60.5.p223

Bruce, D. J. (1958). The effects of listeners' anticipation on the intelligibility of heard speech. *Lang. Speech* 1, 79–97.

Bruner, J. S., and Potter, M. C. (1964). Interference in visual recognition. *Science* 144, 424–425. doi: 10.1126/science.144.3617.424

Caplan, D., and Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behav. Brain Sci.* 22, 77–126. doi: 10.1017/S0140525X99001788

Carpenter, P. A., Miyake, A., and Just, M. A. (1994). "Working memory constraints in comprehension: evidence from individual differences, aphasia, and aging," in ed. *Handbook of Psycholinguistics*, ed. M. Gernsbacher (San Diego, CA: Academic Press), 1075–1122.

Christianson, K., Williams, C. C., Zacks, R. T., and Ferreira, S. (2006). Younger and older adults' "good enough" interpretations of garden path sentences. *Discourse Process.* 42, 205–238. doi: 10.1207/s15326950dp4202_6

Chun, M. M., Golomb, J. D., and Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annu. Rev. Psychol.* 62, 73–101. doi: 10.1146/annurev.psych.093008.100427

Cohen, G., and Faulkner, D. (1983). Word recognition: age differences in contextual facilitation effects. *Br. J. Psychol.* 74, 239–251. doi: 10.1111/j.2044-8295.1983.tb01860.x

Cowan, N. (1999). "An embedded-process model of working memory," in *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, eds A. Miyake and P. Shah (Cambridge: Cambridge University Press), 62–101.

Cowan, N. (2005). "Working memory capacity limits in a theoretical context," in *Human Learning and Memory: Advances in Theory and Application: The 4th Tsukuba International Conference on Memory*, eds C. Izawa and N. Ohta (Mahwah, NJ: Erlbaum), 155–175.

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learning Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546

de Fockert, W., de, Rees, G., Frith, C. D., and Lavie, N. (2001). The role of working memory in visual selective attention. *Science* 291, 1803–1806. doi: 10.1126/science.1056496

Dickinson, C. V. M., and Rabbitt, P. M. A. (1991). Simulated visual impairment: effects on text comprehension and reading speed. *Clin. Vis. Sci.* 6, 301–308.

Engle, R. (2002). Working memory as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1111/1467-8721.00160

Engle, R. W., and Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychol. Learn. Motiv.* 44, 145–199. doi: 10.1016/S0079-7421(03)44005-X

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, R. A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331. doi: 10.1037/0096-3445.128.3.309

Fabry, D. (2011). Jim Jerger by the letters. *Audiol. Today* 19–29.

Fallon, M., Peelle, J. E., and Wingfield, A. (2006). Spoken sentence processing in young and older adults modulated by task demands: evidence from self-paced listening. *J. Genrontol. Psychol. Sci.* 61B, P310–P317. doi: 10.1093/geronb/61.1.p10

Feld, J., and Sommers, M. S. (2009). Lipreading, processing speed, and working memory in younger and older adults. *J. Speech Lang. Hear. Res.* 52, 1555–1565. doi: 10.1044/1092-4388(2009/08-0137)

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cogn. Psychol.* 47, 164–203. doi: 10.1016/S0010-0285(03)00005-7

Ferreira, F., Bailey, K. G. D., and Ferraro, V. (2002). Good-enough representations in language comprehension. *Curr. Dir. Psychol. Sci.* 11, 11–15. doi: 10.1111/1467-8721.00158

Ferreira, F., and Patson, N. (2007). The 'good enough' approach to language comprehension. *Lang. Linguist. Compass* 1, 71–83. doi: 10.1111/j.1749-818X.2007.00007.x

Fillenbaum, S. (1974). Pragmatic normalization: further results for some conjunctive and disjunctive sentences. *J. Exp. Psychol.* 102, 574–578. doi: 10.1037/h0036092

Fisk, J. E., and Sharp, C. A. (2004). Age-related impairment in executive functioning: updating, inhibition, shifting, and access. *J. Clin. Exp. Neuropsychol.* 26, 874–890. doi: 10.1080/13803390490510680

Flores d'Arcais, G. B. (1987). "Automatic processes in language comprehension," in *Perspectives in Cognitive Neuropsychology*, eds G. Denes, C. Semenza, P. Bisiacchi, and E. Andreewsky (Hillsdale, NJ: Erlbaum), 91–114.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Forster, K. J. (1976). "Accessing the mental lexicon," in *New Approaches to Language Mechanisms*, eds R. J. Wales and E. Walker (Amsterdam: North-Holland).

Forster, K. J. (1981). Priming and the effects of sentence and lexical contexts on naming time: evidence for autonomous lexical processing. *Q. J. Exp. Psychol.* 33, 465–495. doi: 10.1080/14640748108400804

Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6:347.doi: 10.3389/fnagi.2014.00347

Goldman-Eisler, F. G. (1968). *Psycholinguistics: Experiments in Spontaneous Speech*. New York, NY: Academic Press.

Gordon-Salant, S., and Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *J. Speech Lang. Hear. Res.* 40, 423–431. doi: 10.1044/jslhr.4002.423

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Percept. Psychophys.* 28, 267–283. doi: 10.3758/BF03204386

Grosjean, F. (1985). The recognition of words after their acoustic offset: evidence and implications. *Percept. Psychophys.* 38, 299–310. doi: 10.3758/BF03207159

Grosjean, F. (1996). Gating. *Lang. Cogn. Process.* 11, 597–604. doi: 10.1080/016909696386999

Hasher, L., Stoltzfus, E. R., Zacks, R. T., and Rypma, B. R. (1991). Age and inhibition. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 163–169. doi: 10.1037/0278-7393.17.1.163

Hasher, L., and Zacks, R. T. (1988). "Working memory, comprehension, and aging: a review and a new view," in *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 22, ed. G. H. Bower (San Diego, CA: Academic Press), 193–225.

Heald, S. L. M., and Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Front. Syst. Neurosci.* 8:35. doi: 10.3389/fnsys.2014.00035

Holcomb, P. J., and Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. *Lang. Cogn. Process.* 5, 281–312. doi: 10.1080/01690969008407065

Howes, D. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *J. Acoust. Soc. Am.* 29, 296–305. doi: 10.1121/1.1908862

Humes, L. E. (1996). Speech understanding in the elderly. *J. Am. Acad. Audiol.* 7, 161–167.

Hunnicutt, S. (1985). Intelligibility versus redundancy—conditions of dependency. *Lang. Speech* 28, 47–56.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149. doi: 10.1037/0033-295X.99.1.122

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kantowitz, B. H., and Knight, J. L. (1976). On experimenter limited processes. *Psychol. Rev.* 83, 502–507. doi: 10.1037/0033-295X.83.6.502

LaBerge, D., and Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cogn. Psychol.* 6, 293–323. doi: 10.1016/0010-0285(74)90015-2

Lahar, C. J., Tun, P. A., and Wingfield, A. (2004). Sentence-final word completion norms for young, middle-aged, and older adults. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 59, P7–P10. doi: 10.1093/geronb/59.1.p7

Lash, A., Rogers, C. S., Zoller, A., and Wingfield, A. (2013). Expectation and entropy in spoken word recognition: effects of age and hearing acuity. *Exp. Aging Res.* 39, 235–253. doi: 10.1080/0361073X.2013.779175

Lash, A., and Wingfield, A. (2014). A Bruner-Potter effect in audition? Spoken word recognition in adult aging. *Psychol. Aging* 29, 907–912. doi: 10.1037/a0037829

Lavie, N., Hirst, A., de Fockert, J. W., and Viding, E. (2004). Load theory of selective attention and cognitive control. *J. Exp. Psychol. Gen.* 133, 339–354. doi: 10.1037/0096-3445.133.3.339

Lethbridge-Ceijku, M., Schiller, J. S., and Brernadel, L. (2004). Summary health statistics for U.S. adults: National Health Interview Survey. *Vital Health Stat.* 10, 1–151.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.* 74, 431–461. doi: 10.1037/h0020279

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Lang. Speech* 6, 172–187.

Lindblom, B., Brownlee, S., Davis, B., and Moon, S. J. (1992). Speech transforms. *Speech Commun.* 11, 357–368. doi: 10.1016/0167-6393(92)90041-5

Lindfield, K. C., Wingfield, A., and Goodglass, H. (1999). The contribution of prosody to spoken word recognition. *Appl. Psycholinguist.* 20, 395–405. doi: 10.1017/S0142716499003045

Logie, R. H. (2011). The functional organization and capacity of working memory. *Curr. Dir. Psychol. Sci.* 20, 240–245. doi: 10.1177/0963721411415340

Logie, R. H., Della Sala, S., Laiacona, M., Chalmers, P., and Wynn, V. (1996). Group aggregates and individual reliability: the case of verbal short-term memory. *Mem. Cogn.* 24, 305–321. doi: 10.3758/BF03213295

Luce, P. A., and Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear Hear.* 19, 1–36. doi: 10.1097/00003446-199802000-00001

Luo, C. R., and Snodgrass, J. G. (1994). Competitive activation model of perceptual interference in picture and word identification. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 50. doi: 10.1037/0096-1523.20.1.50

Madden, D. J. (1988). Adult age differences in the effects of sentence context and stimulus degradation during visual word recognition. *Psychol. Aging* 3, 167–172. doi: 10.1037/0882-7974.3.2.167

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition* 25, 71–102. doi: 10.1016/0010-0277(87)90005-9

Marslen-Wilson, W. D., and Zwitserlood, P. (1989). Accessing spoken words: the importance of word onsets. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 576–585. doi: 10.1037/0096-1523.15.3.576

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., and Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology* 24, 222–243. doi: 10.1037/a0017619

McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech recognition. *Cogn. Psychol.* 18, 1–86. doi: 10.1016/0010-0285(86)90015-0

Miller, P., and Wingfield, A. (2010). Distinct effects of perceptual quality on auditory word recognition, memory formation and recall in a neural model of sequential memory. *Front. Syst. Neurosci.* 4:14. doi: 10.3389/fnsys.2010.00014

Mishra, S., Lunnrer, T., Stenfelt, S., Rönnberg, J., and Rudner, M. (2013). Seeing the talker's face supports executive processing of speech in steady state noise. *Front. Syst. Neurosci.* 7:96. doi: 10.3389/fnsys.2013.00096/full

Morton, J. (1964a). A preliminary functional model for language behavior. *Int. Audiol.* 3, 1–9. doi: 10.3109/05384916409074089

Morton, J. (1964b). A model for continuous language behavior. *Lang. Speech* 7, 450–470.

Morton, J. (1969). Interaction of information in word recognition. *Psychol. Rev.* 76, 165–178. doi: 10.1037/h0027366

Morton, J. (1979). "Facilitation in word recognition: experiments causing change in the logogen model," in *Processing Visual Language*, eds P. A. Kolers, M. E. Wrolstad, and H. Bouma (New York: Plenum Press), 259–268.

Mullennix, J. W., Pisoni, D. B., and Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *J. Acous. Soc. Am.* 85, 365–378. doi: 10.1121/1.397688

Murphy, D. R., Craik, F. I. M., Li, K. Z. H., and Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* 15, 323–334. doi: 10.1037/0882-7974.15.2.323

Newby, H. A. (1958). *Audiology: Principles and Practice*. New York, NY: Appleton-Century-Crofts.

Norman, D. A., and Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cogn. Psychol.* 7, 44–64. doi: 10.1016/0010-0285(75)90004-3

Nusbaum, H., and Magnuson, J. (1997). "Talker normalization: phonetic constancy as a cognitive process," in *Talker Variability and Speech Processing*, K. A. Johnson and J. W. Mullennix (New York, NY: Academic Press), 109–132.

Pashler, H., Johnston, J. C., and Ruthruff, E. (2001). Attention and performance. *Annu. Rev. Psychol.* 52, 629–651. doi: 10.1146/annurev.psych.52.1.629

Pichora-Fuller, M. K. (2003). Cognitive aging and auditory information processing. *Int. J. Audiol.* 42, 2S26–2S32. doi: 10.3109/14992020309074641

Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *J. Acoust. Soc. Am.* 97, 593–607. doi: 10.1121/1.412282

Piquado, T., Benichov, J. I., Brownell, H., and Wingfied, A. (2012). The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening. *Int. J. Audiol.* 51, 576–583. doi: 10.3109/14992027.2012.684403

Piquado, T., Cousins, K. A. Q., Wingfield, A., and Miller, P. (2010). Effects of degraded sensory input on memory for speech: behavioral data and a test of biologically constrained computational models. *Brain Res.* 1365, 48–65. doi: 10.1016/j.brainres.2010.09.070

Pollack, I., and Pickett, J. M. (1963). The intelligibility of excerpts from conversation. *Lang. Speech* 6, 165–171.

Posner, M. I., and Snyder, C. R. R. (1975). "Attention and cognitive control," in *Information Processing and Cognition: The Loyola Symposium*, ed. R. Solso (Hillsdale, NJ: Erlbaum), 205–223.

Potter, M. C. (1993). Very short-term conceptual memory. *Mem. Cogn.* 21, 156–161. doi: 10.3758/BF03202727

Potter, M. C., and Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *J. Mem. Lang.* 29, 633–654. doi: 10.1016/0749-596X(90)90042-X

Rabbitt, P. (1968). Channel capacity, intelligibility, and immediate memory. *Q. J. Exp. Psychol.* 20, 241–248. doi: 10.1080/14640746808400158

Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol.* 476, 167–176. doi: 10.3109/00016489109127274

Rogers, C. S., Jacoby, L. L., and Sommers, M. S. (2012). Frequent false hearing by older adults: the role of age differences in metacognition. *Psychol. Aging* 27, 33–45. doi: 10.1037/a0026231

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int. J. Audiol.* 42, S68–S76. doi: 10.3109/14992020309074626

Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise Health* 12, 263.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013. 00031

Rönnberg, J., Lyxell, B., Arlinger, S., and Kinnefors, C. (1989). Visual evoked potentials: relation to adult speechreading and cognitive function. *J. Speech Hear. Lang. Res.* 32, 725–735. doi: 10.1044/jshr.3204.725

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S171–S177. doi: 10.1080/14992020802301167

Rudner, M., and Lunner, T. (2013). Cognitive spare capacity as a window on hearing aid benefit. *Semin. Hear.* 34, 298–307. doi: 10.1055/s-0033-1356642

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology* 8, 535–543. doi: 10.1037/0894-4105.8.4.535

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403

Salthouse, T. A., Atkinson, T. M., and Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *J. Exp. Psychol. Gen.* 132, 566–594. doi: 10.1037/0096-3445.132.4.566

Samuels, A. G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychol. Sci.* 12, 348–351. doi: 10.1111/1467-9280.00364

Sanford, A. J., and Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends Cogn. Sci.* 6, 382–386. doi: 10.1016/S1364-6613(02)01958-7

Schacter, D. L. (1987). Implicit memory: history and current status. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 501–518. doi: 10.1037/0278-7393.13.3.501

Schneider, B. A., and Pichora-Fuller, M. K. (2000). "Implications of perceptual deterioration for cognitive aging research," in *Handbook of Aging and Cognition*, 2nd Edn, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ.: Erlbaum), 155–219.

Schneider, W., and Chen, J. M. (2003). Controlled and automatic processing: behavior, theory, and biological mechanisms. *Cogn. Sci.* 27, 525–559. doi: 10.1207/s15516709cog2703_8

Schoof, T., and Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Front. Aging Neurosci.* 6:307. doi: 10.3389/fnagi.2014.00307

Seidenberg, M. S., Tannenhaus, M. K., Leiman, J. M., and Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing. *Cogn. Psychol.* 14, 489–537. doi: 10.1016/0010-0285(82)90017-2

Seidenberg, M. S., Waters, G. S., Sanders, M., and Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Mem. Cogn.* 12, 315–328. doi: 10.3758/BF03198291

Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol. Rev.* 84, 155–171. doi: 10.1037/0033-295X.84.2.127

Snodgrass, J. G., and Hirshman, E. (1991). Theoretical explorations of the Bruner-Potter 1964 interference effect. *J. Mem. Lang.* 30, 273–293. doi: 10.1016/0749-596X(91)90037-K

Sommers, M. S., and Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: the interaction of lexical competition and semantic context. *Psychol. Aging* 14, 458–472. doi: 10.1037/0882-7974.14.3.458

Sommers, M. S., Spehar, B., and Tye-Murray, N. (2005). Auditory visual speech perception and visual enhancement in normal-hearing younger and older adults. *Ear Hear.* 26, 263–275. doi: 10.1097/00003446-200506000-00003

Sörqvist, P., Stenfelt, S., and Rönnberg, J. (2012). Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: toward a unified view of attention. *J. Cogn. Neurosci.* 24, 2147–2154. doi: 10.1162/jocn_a_00275

Squire, L. R. (1994). "Declarative and nondeclarative memory: multiple brain systems supporting learning and memory," in *Memory Systems*, eds D. L. Schacter and E. Tulving (Cambridge, MA: MIT Press), 203–232.

Stanovich, K. E., and West, R. F. (1983). On priming by sentence context. *J. Exp. Psychol. Gen.* 112, 1–36. doi: 10.1037/0096-3445.112.1.1

Stine, E. A. L., Wingfield, A., and Myers, S. D. (1990). Age differences in processing information from television news: the effects of bisensory augmentation. *J. Gerontol. Psychol. Sci.* 45, 1–8. doi: 10.1093/geronj/45.1.P1

Stromswold, K., Caplan, D., Alpert, N., and Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* 52, 452–473. doi: 10.1006/brln.1996.0024

Sumby, W., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215. doi: 10.1121/1.1907309

Surprenant, A. M. (1999). The effect of noise on memory for spoken syllables. *Int. J. Psychol.* 34, 328–333. doi: 10.1080/002075999399648

Surprenant, A. M. (2007). Effects of noise on identification and serial recall of nonsense syllables in older and younger adults. *Aging Neuropsychol. Cogn.* 14, 126–143. doi: 10.1080/13825580701217710

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *J. Verbal Learn. Verbal Behav.* 18, 645–659. doi: 10.1016/S0022-5371(79)90355-4

Tulving, E. (1972). "Episodic and semantic memory," in *Organization of Memory*, eds E. Tulving and W. Donaldson (New York, NY: Academic Press), 381–403.

Tyler, L. K. (1984). The structure of the initial cohort: evidence from gating. *Percept. Psychophys.* 36, 417–427. doi: 10.3758/BF03207496

van Rooij, J. C. G. M., and Plomp, R. (1992). Auditive and cognitive factors in speech perception by elderly listeners: III. Additional data and final discussion. *J. Acoust. Soc. Am.* 91, 1028–1033. doi: 10.1121/1.402628

Wagenmakers, E-J., Ratcliff, R., Gomez, P., and McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *J. Mem. Lang.* 58, 140–159. doi: 10.1016/j.jml.2007.04.006

Waters, G., and Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Q. J. Exp. Psychol.* 49A, 51–79. doi: 10.1080/713755607

Wayland, S. C., Wingfield, A., and Goodglass, H. (1989). Recognition of isolated words: the dynamics of cohort reduction. *Appl. Psycholinguist.* 10, 475–487. doi: 10.1017/S0142716400009048

Wingfield, A., Aberdeen, J. S., and Stine, E. A. L. (1991). Word onset gating and linguistic context in spoken word recognition by young and elderly adults. *J. Gerontol. Psychol. Sci.* 46, P127–P129. doi: 10.1093/geronj/46.3.p127

Wingfield, A., Alexander, A. H., and Cavigelli, S. (1994). Does memory constrain utilization of top-down information in spoken word recognition? Evidence from normal aging. *Lang. Speech* 37, 221–235.

Wingfield, A., Goodglass, H., and Lindfield, K. C. (1997). Word recognition from acoustic onsets and acoustic offsets: effects of cohort size and syllabic stress. *Appl. Psycholinguist.* 18, 85–100. doi: 10.1017/S0142716400009887

Wingfield, A., Goodglass, H., and Smith, K. L. (1990). Effects of word-onset cuing on picture naming in aphasia: a reconsideration. *Brain Lang.* 39, 373–390. doi: 10.1016/0093-934X(90)90146-8

Wingfield, A., and Lindfield, K. C. (1995). Multiple memory systems in the processing of speech: evidence from aging. *Exp. Aging Res.* 21, 101–121. doi: 10.1080/03610739508254272

Wingfield, A., McCoy, S. L., Peelle, J. E., Tun, P. A., and Cox, L. C. (2006). Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. *J. Am. Acad. Audiol.* 17, 487–497. doi: 10.3766/jaaa.17.7.4

Wingfield, A., Stine, E. L., Lahar, C. J., and Aberdeen, J. S. (1988). Does the capacity of working memory change with age? *Exp. Aging Res.* 14, 103–107. doi: 10.1080/03610738808259731

Wingfield, A., and Stine-Morrow, E. A. L. (2000). "Language and speech," in *Handbook of Aging and Cognition*, 2nd Edn, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Erlbaum), 359–416.

Wingfield, A., and Tun, P. A. (2001). Spoken language comprehension in older adults: interactions between sensory and cognitive change in normal aging. *Semin. Hear.* 22, 287–301. doi: 10.1055/s-2001-15632

Wingfield, A., Tun, P. A., Koh, C. K., and Rosen, M. J. (1999). Regaining lost time: adult aging and the effect of time restoration on recall of time-compressed speech. *Psychol. Aging* 14, 380–389. doi: 10.1037/0882-7974.14.3.380

Wingfield, A., Tun, P. A., and McCoy, S. L. (2005). Hearing loss in older adulthood: what it is and how it interacts with cognitive performance. *Curr. Dir. Psychol. Sci.* 14, 144–148. doi: 10.1111/j.0963-7214.2005.00356.x

Wingfield, A., Waters, G. S., and Tun, P. A. (1998). "Does working memory work in language comprehension?: evidence from behavioral neuroscience," in *The Other Side of the Error Term: Aging and Development as Model Systems in Cognitive Neuroscience*, ed. N. Raz (Amsterdam: Elsevier), 319–393.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Wingfield, Amichetti and Lash. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read, for greatest visibility

COLLABORATIVE PEER-REVIEW
Designed to be rigorous – yet also collaborative, fair and constructive

FAST PUBLICATION
Average 85 days from submission to publication (across all journals)

COPYRIGHT TO AUTHORS
No limit to article distribution and re-use

TRANSPARENT
Editors and reviewers acknowledged by name on published articles

SUPPORT
By our Swiss-based editorial team

IMPACT METRICS
Advanced metrics track your article's impact

GLOBAL SPREAD
5'100'000+ monthly article views and downloads

LOOP RESEARCH NETWORK
Our network increases readership for your article

Frontiers
EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on