



OPEN ACCESS

EDITED BY

Eva Wittenberg,
Central European University, Hungary

REVIEWED BY

Susan E. Teubner-Rhodes,
Auburn University, United States
Giuseppina Turco,
Délégation Paris-Villejuif-01 (CNRS), France

*CORRESPONDENCE

Akshay R. Maggu
✉ akshay.maggu@uconn.edu

RECEIVED 26 September 2025

REVISED 07 November 2025

ACCEPTED 17 November 2025

PUBLISHED 02 December 2025

CITATION

Sharma B, Li Q and Maggu AR (2025) Unreal words, real competition: Mandarin recognition slows for syllable-matched tonal gaps. *Front. Lang. Sci.* 4:1714072. doi: 10.3389/flang.2025.1714072

COPYRIGHT

© 2025 Sharma, Li and Maggu. 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) and the copyright owner(s) 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.

Unreal words, real competition: Mandarin recognition slows for syllable-matched tonal gaps

Bhamini Sharma¹, Qun Li² and Akshay R. Maggu^{3,4*}

¹Department of Communication Arts, Sciences, and Disorders, City University of New York – Brooklyn College, Brooklyn, NY, United States, ²Department of Audiology and Speech-Language Pathology, Department of Otorhinolaryngology Head and Neck Surgery, West China Hospital of Sichuan University, Chengdu, China, ³Department of Speech, Language, and Hearing Sciences, University of Connecticut, Storrs, CT, United States, ⁴Institute for Brain and Cognitive Sciences, University of Connecticut, Storrs, CT, United States

Word recognition in tone languages like Mandarin is influenced not only by phonological structure but also by lexical tone. Prior research using auditory lexical decision tasks has shown that real monosyllables are generally processed more quickly and accurately than tonal gaps i.e., impossible syllable-tone combinations. However, these studies often did not control for syllable overlap between tonal gaps and real monosyllables, potentially underestimating lexical competition effects. The present study addressed this gap by contrasting real monosyllables, syllable-matched tonal gaps, and syllable-unmatched tonal gaps in a controlled auditory lexical decision task with 54 native Mandarin speakers. Results revealed that reaction times were significantly faster and accuracy higher for real monosyllables compared to both syllable-matched tonal gaps and syllable-unmatched tonal gaps. More importantly, syllable-matched tonal gaps elicited slower reaction times than syllable-unmatched tonal gaps, indicating increased lexical competition when tonal gaps share the same syllable as real monosyllables. These findings emphasize the critical role of phonological similarity and lexical competition in Mandarin word recognition. By controlling syllabic overlap, this study improves upon previous methodologies and offers a clearer assessment of auditory word processing in tone languages.

KEYWORDS

Mandarin word recognition, auditory lexical decision, tonal gaps, non-word processing, reaction time

Introduction

Word recognition is a complex multifaceted process that involves the decoding and identification of meaning. Research methodology in understanding word recognition (Slowiaczek and Pisoni, 1986; Milberg et al., 1988; Goldinger, 1996; Luce and Pisoni, 1998; Vitevitch and Luce, 1999; Ernestus and Cutler, 2015; Ferrand et al., 2018; Nenadić and Tucker, 2020; Nenadić et al., 2023) often involves contrasting the processing of real words and non-words in an auditory lexical decision task (ALDT). This word recognition process, in tone languages such as Mandarin, is further complicated by the presence of lexical tones that signal meaning. In order to study word recognition in Mandarin, existing studies (Wang et al., 2012; Wiener and Turnbull, 2016; Yao and Sharma, 2017; Neergaard, 2018; Sharma, 2020; Chang and Hsieh, 2022; Gong et al., 2024) have utilized tonal gaps—combinations of syllables and tones that are not possible in the language—along with real monosyllables (RW). However, ALDT in the existing studies in Mandarin (Wang et al., 2012; Wiener and Turnbull, 2016; Yao and Sharma, 2017; Neergaard, 2018; Sharma, 2020)

involved tonal gaps and RW that did not match in terms of syllable, causing a potentially reduced competition in delineation of the two types of stimuli. For example, *ta2* which is a tonal gap did not have a corresponding real monosyllable *ta1* in the previous experiments. To get a more fine-grained understanding into Mandarin word recognition, in the current study, we used tonal gaps that matched the RW on syllables (e.g., *ta2* tonal gap and *ta1* real monosyllable that match on syllable *ta*), and tonal gaps that did not match the RW on syllables (e.g., *le2* is a tonal gap for which no corresponding real monosyllable sharing the syllable *le* was included in the present stimulus set).

Mandarin is a tone language where monosyllables serve as the fundamental units of the lexicon, with each monosyllable corresponding to a Chinese character. A monosyllable consists of an onset, a rime, and a lexical tone. The onset refers to the consonant, while the rime includes a vowel or diphthong, with or without a coda, and is combined with one of the four lexical tones in Mandarin. When tones are excluded, there are roughly 400 possible syllables, but when tones are considered, the total increases to around 1,300 syllables which correspond to at least one Chinese character. This discrepancy highlights that not all potential combinations of syllables and tones are possible resulting in *tonal gaps*. Tonal gaps are impossible combinations of syllables and tones, where a particular syllable may appear with other tones but not with a specific tone. For example, while monosyllables such as *ta1* (他), *ta3* (塔), and *ta4* (踏) exist in Mandarin, *ta2* does not, making *ta2* a tonal gap.

Previous research on word recognition in Mandarin using ALDT has shown varying results regarding the processing of RW and tonal gaps. Some studies have found differences in accuracy and reaction times (RT) between these two types of stimuli (Wang et al., 2012; Wiener and Turnbull, 2016; Yao and Sharma, 2017; Neergaard, 2018; Sharma, 2020), while others have not specifically analyzed the pseudosyllables (Wang et al., 2012). Overall, RW are recognized significantly faster compared to tonal gaps (Wang et al., 2012; Wiener and Turnbull, 2016; Yao and Sharma, 2017; Neergaard, 2018; Sharma, 2020), and there is some evidence suggesting they are also recognized more accurately (Yao and Sharma, 2017; Sharma, 2020).

It is important to note that these studies have typically included not only RW and tonal gaps but also syllable gaps, which are combinations of onset, rime, and tone that do not exist in any form in Mandarin, as part of their experimental designs (Wiener and Turnbull, 2016; Yao and Sharma, 2017; Neergaard, 2018; Sharma, 2020). Additionally, some studies have not clearly indicated whether tonal gaps were included in their tasks. A major limitation in this body of research is that none of these studies have matched tonal gaps with its corresponding RW in terms of syllable, within the experiment (e.g., for *ta1* real monosyllable, there is no *ta2* tonal gap included in the previous studies' experiments). This omission may potentially dilute the competition between tonal gaps and RW in ALDT, potentially affecting the validity of the findings. The present study aims to address this gap by including matched RW corresponding to the tonal gaps, thereby providing a more accurate reflection of how these stimuli may compete during Mandarin word recognition.

In the current study, to gain a more fine-grained understanding of Mandarin word recognition, we used stimuli that included tonal gaps matching the RW in syllable [hereafter, *syllable matched tonal gaps* (MTG)], tonal gaps not matching the RW in syllable [hereafter, *syllable unmatched tonal gaps* (UTG)], and real monosyllables, in an ALDT paradigm with native speakers of Mandarin as subjects. We hypothesized that lexical competition would be influenced by the degree of syllable overlap among stimuli. Specifically, we expected that RW would elicit faster, and more accurate responses compared to both types of tonal gaps. Furthermore, we predicted that MTG would result in longer reaction times and slightly reduced accuracy relative to UTG, reflecting stronger lexical competition from their real-word counterparts.

Method

Participants

A total of 59 participants (12 males, 47 females; mean age = 22.10 years, SD = 1.88) took part in the study. All participants were recruited from the Sichuan University and were native Mandarin speakers, having been born and raised in mainland China. The research was approved by the Research Ethics Committee at the West China Hospital of Sichuan University (# 2023–2376). Informed written consent was obtained from all participants. The participants reported no history of speech or hearing impairments. All the participants were screened for hearing in both ears from 250–8,000 Hz and only those who had ≤ 20 dB HL on all tested frequencies were enrolled to take part in the experiment. As a result, five participants were removed, and we went ahead with 54 participants.

Materials

The experiment used the existing material (Yao and Sharma, 2017; Sharma, 2020) containing 216 monosyllabic stimuli, divided into two categories: 108 RW and 108 pseudosyllables (see Table 1). The pseudosyllables were accidental gaps due to lexical tone in Mandarin, also known as tonal gaps. Of the 108 pseudosyllables,

TABLE 1 Distribution of stimuli by type, number of items, and representative example stimuli used in the experiment.

Type of stimuli	Number of items	Example
Real monosyllables (RW)	30	ta1
Real monosyllables balanced for lexical tone with RW	30	fang1
Syllable matched tonal gaps (MTG)	30	ta2
Real monosyllables balanced for lexical tone with MTG	30	fang2
Syllable unmatched tonal gaps (UTG)	78	xin3
Real monosyllable fillers	18	kan4

30 were MTG, while the remaining 78 served as UTG. For the 108 real monosyllables, 30 were RW, 30 were real monosyllables balanced for lexical tone with the RW, 30 were real monosyllables balanced for lexical tone with the MTG, and the final 18 were real monosyllable fillers. The distribution of stimuli across subcategories (see Table 1) was necessary to meet the requirements of ALDT, in which 50% of the stimuli must be legal (real) syllables and 50% illegal (pseudosyllables) to maintain task balance and control decision-making processes. Although six stimulus categories were created to achieve tonal and lexical balance, only the 30 RW, 30 MTG, and 78 UTG were included in the main analyses; the remaining real-monosyllable sets served to ensure tonal balance but were excluded from statistical comparisons. All stimuli were recorded by a female native speaker of Mandarin. The stimuli had a mean duration of 612 ms ($SD = 96$ ms), and a one-way ANOVA revealed no significant differences in duration across the RW, MTG, and UTG conditions, $F_{(2, 135)} = 0.14$, $p = 0.874$. Given the absence of significant variation, reaction times were analyzed without applying any duration-based correction.

Procedure

ALDT was carried out using E-Prime 3.0 (Psychology Software Tools, Inc, 2016). Participants were presented with a randomized list of 216 stimuli, each presented twice throughout the experiment. The experiment was run on a ThinkPad X1 Carbon Gen 10 LTE1 laptop, connected to Sennheiser HD 280 Pro studio monitoring headphones. The experiment took place in a soundproof room to minimize environmental noise interference.

Before the experiment began, participants were seated comfortably in front of the laptop. Each trial began with a fixation cross displayed at the center of the screen for 500 ms, followed by the auditory stimulus. A 4,000 ms response window began at stimulus onset, allowing approximately 3.3 s after the average stimulus duration of 612 ms. This duration provided participants with sufficient time to fully process the auditory stimulus and determine whether it could be associated with a legal Chinese character before executing a response. Participants were instructed to listen carefully to each single-syllable Mandarin sound and determine, as quickly and accurately as possible, whether it corresponded to an actual word in Mandarin. If the sound could be linked to at least one existing Chinese character (i.e., a real monosyllable), they pressed the “A” key. If the sound could not be associated with any known character (i.e., a pseudo-syllable or tonal gap), they pressed the “L” key. The keys were counterbalanced across participants. If the participant did not respond within the 4,000 ms time frame, the trial automatically advanced to the next stimulus.

RT and accuracy were recorded. RT were recorded from the onset of the stimulus until the participant’s response. Before beginning the main experiment, participants completed a practice session with 12 items, during which feedback was provided. None of the stimuli in the practice session were used in the main experiment. The entire experimental session lasted no more than 20 min.

Results

The data from 54 native young adult speakers of Mandarin were analyzed for RT and accuracy. RT analyses were conducted exclusively on correct trials. Broadly, the results of the study indicated that RT were shortest for RW (mean = 1,044.14 ms; $SD = 155.11$ ms) compared to tonal gaps (mean = 1,477.62 ms; $SD = 391.16$ ms), and accuracy was highest for RW (mean = 0.90; $SD = 0.07$) when compared to tonal gaps (mean = 0.83; $SD = 0.16$). Within the tonal gaps category, the MTG resulted in the longer RT as compared to UTG (i.e., 1,516.33 ms for MTG vs. 1,438.89 ms for UTG). Regarding accuracy, the highest performance was observed for RW followed by accuracy for MTG followed by UTG (i.e., $RW = 0.90 > MTG = 0.85 > UTG = 0.84$). Conditional accuracy functions (CAFs), computed across five RT quintiles, revealed consistently higher accuracy for MTG compared to UTG, with no crossover of curves. Based on the CAFs, MTG responses remained more accurate than UTG responses even at slower RTs, potentially due to increased lexical competition rather than a speed–accuracy trade-off (Supplementary Figure S1).

To compare RW with MTG and UTG (i.e., the tonal gaps examined in the previous studies), we conducted inferential statistical analysis using RT and accuracy as dependent measures. Repeated measures ANOVA revealed a significant difference in RT across conditions [$F_{(1.63)} = 73.94$, $p < 0.001$]. *Post-hoc* pairwise comparisons using the Holm correction revealed significant differences between RW and MTG ($p < 0.0001$), RW and UTG ($p < 0.0001$), and between MTG and UTG ($p = 0.02$; Figure 1A).

For accuracy, as the assumption of normality was not fully met, as one of the conditions (RW) showed a significant deviation from normality (RW: $W = 0.928$, $p = 0.0038$; MTG: $W = 0.961$, $p = 0.0781$; UTG: $W = 0.980$, $p = 0.518$), a Kruskal-Wallis test was conducted, which showed a significant difference between the three stimuli conditions [$H(2) = 11.32$, $p = 0.003$]. Pairwise comparisons using the Wilcoxon rank-sum test revealed a significant difference between RW and UTG ($p = 0.002$; Figure 1B).

Discussion

The present study aimed to provide a more nuanced understanding of Mandarin word recognition by controlling for syllabic overlap between RW and tonal gaps—an important methodological refinement over previous studies. Our findings revealed significantly faster RT and higher accuracy for RW compared to tonal gaps (both MTG and UTG). More importantly, we found significantly slower RT for MTG compared to UTG supporting our hypothesis that when tonal gaps share the same syllable structure as the real monosyllable, lexical competition increases. Although accuracy generally followed a similar overall pattern to reaction times ($RW > MTG > UTG$), the relationship between MTG and UTG revealed an important dissociation. More specifically, CAFs indicated that MTG responses remained more accurate than UTG responses across slower reaction times, consistent with increased lexical competition rather than a speed–accuracy trade-off.

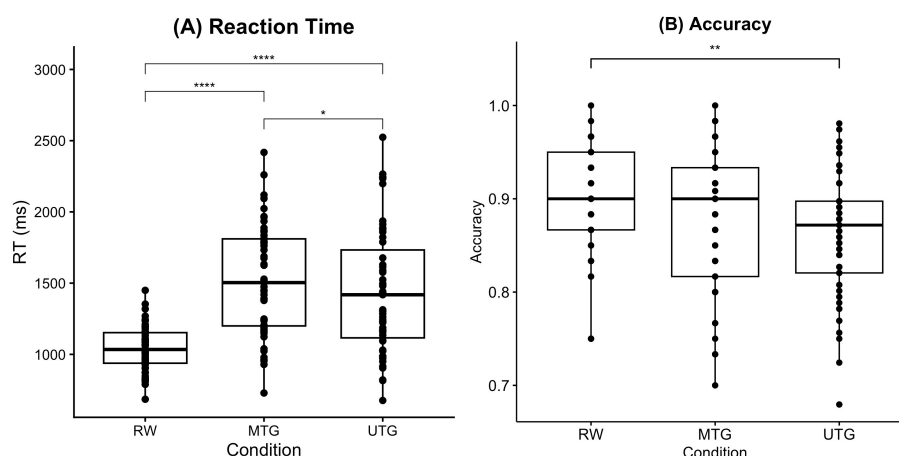


FIGURE 1

Comparison of real monosyllables (RW), syllable matched tonal gaps (MTG), and syllable unmatched tonal gaps (UTG) on (A) reaction time (RT) and (B) accuracy. * $p < 0.05$; *** $p < 0.01$; **** $p < 0.0001$.

These findings are broadly consistent with previous research on Mandarin word recognition using ALDT, which has shown mixed results regarding the processing of RW and tonal gaps. Some studies have reported significant differences in RTs and accuracy between RW and tonal gaps (Wang et al., 2012; Wiener and Turnbull, 2016; Yao and Sharma, 2017; Neergaard, 2018; Sharma, 2020), while others have not specifically analyzed pseudosyllables (Wang et al., 2012). Overall, real words are recognized significantly faster than tonal gaps (Wang et al., 2012; Wiener and Turnbull, 2016; Yao and Sharma, 2017; Neergaard, 2018; Sharma, 2020), and some evidence also suggests higher accuracy for RW (Yao and Sharma, 2017; Sharma, 2020). However, unlike prior research, our study employed a more rigorously controlled design where tonal gaps and RW were matched on syllable structure and found that there was a significant difference in RT for tonal gaps matched for syllables (i.e., MTG) with the RW and tonal gaps not matched for syllables (i.e., UTG) with the RW. This methodological improvement allows for a more accurate examination of lexical competition. In earlier studies, tonal gaps and RW did not match on syllabic composition, which could have inadvertently decreased the cognitive competition between stimulus types and weakened the validity of the results.

The slower rejection of syllable-matched tonal gaps likely reflects transient activation of specific lexical neighbors that share identical syllabic frames, consistent with neighborhood activation models of spoken word recognition (Luce and Pisoni, 1998; Vitevitch and Luce, 1999). In contrast, syllable-unmatched tonal gaps may elicit broader activation at the phonotactic level. Future studies using a controlled experimental design that systematically manipulates the presence or absence of syllable-matched tonal gaps could more directly test their influence on real-word recognition latencies and further elucidate how contextual factors shape lexical competition.

Our results highlight the importance of fine-grained phonological control in auditory lexical decision tasks, particularly in tone languages like Mandarin, where monosyllables and tonal distinctions carry lexical meaning. The finding that even UTG and MTG elicited significantly different RT supports the robustness of the lexical competition effect under carefully controlled conditions.

In conclusion, this study demonstrated that Mandarin speakers experience greater processing times when encountering syllable-matched tonal gaps compared to unmatched tonal gaps, highlighting the critical role of phonological similarity and lexical competition in tone language perception. The improved experimental control in this study offers a clearer window into these dynamics and paves the way for future research examining tone language processing in populations with language or hearing impairments, as well as in developmental and aging contexts. Additionally, this design can be adapted for cross-linguistic comparisons in other tone or pitch-accent languages, further broadening its applicability.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Research Ethics Committee at the West China Hospital of Sichuan University (# 2023-2376). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

BS: Writing – review & editing, Methodology, Conceptualization, Supervision, Writing – original draft, Investigation, Project administration, Visualization, Resources, Formal analysis. QL: Investigation, Writing – review & editing, Data curation. AM: Investigation, Funding acquisition, Conceptualization, Supervision, Resources, Writing – review & editing, Formal analysis, Project administration, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The article processing charges were covered by AM through their research fund.

Conflict of interest

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have

been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/flang.2025.1714072/full#supplementary-material>

References

- Chang, C.-Y., and Hsieh, F. (2022). Do subsyllabic units play a role in Mandarin spoken word recognition? Evidence from phonotactic processing. *J. Neurolinguist.* 64:101089. doi: 10.1016/j.jneuroling.2022.101089
- Ernestus, M., and Cutler, A. (2015). BALDEY: a database of auditory lexical decisions. *Q. J. Exp. Psychol.* 68, 1469–1488. doi: 10.1080/17470218.2014.984730
- Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., et al. (2018). MEGALEX: a megastudy of visual and auditory word recognition. *Behav. Res.* 50, 1285–1307. doi: 10.3758/s13428-017-0943-1
- Goldinger, S. D. (1996). Auditory lexical decision. *Lang. Cogn. Process.* 11, 559–568. doi: 10.1080/016909696386944
- Gong, S., Zhang, J., and Fiorentino, R. (2024). Phonological well-formedness constraints in mandarin phonotactics: evidence from lexical decision. *Lang. Speech* 67, 676–691. doi: 10.1177/00238309231182363
- 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
- Milberg, W., Blumstein, S., and Dworetzky, B. (1988). Phonological factors in lexical access: evidence from an auditory lexical decision task. *Bull. Psychon. Soc.* 26, 305–308. doi: 10.3758/BF03337665
- Neergaard, K. D. (2018). *Phonological Segmentation Neighborhoods*. Hong Kong: The Hong Kong Polytechnic University.
- Nenadić, F., and Tucker, B. V. (2020). Computational modelling of an auditory lexical decision experiment using jTRACE and TISK. *Lang. Cogn. Neurosci.* 35, 1326–1354. doi: 10.1080/23273798.2020.1764600
- Nenadić, F., Tucker, B. V., and Ten Bosch, L. (2023). Computational modeling of an auditory lexical decision experiment using DIANA. *Lang. Speech* 66, 564–605. doi: 10.1177/00238309221111752
- Psychology Software Tools, Inc. (2016). *E-Prime 3.0*. Available online at: <https://support.psnet.com/> (accessed January 15, 2025).
- Sharma, B. (2020). The effect of phonological neighbors and homophones on spoken word recognition in Mandarin Chinese | PolyU Institutional Research Archive. The Hong Kong Polytechnic University. Available online at: <https://ira.lib.polyu.edu.hk/handle/10397/87412> (Accessed March 27, 2022).
- Slowiczek, L. M., and Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Mem. Cogn.* 14, 230–237. doi: 10.3758/BF03197698
- Vitevitch, M. S., and Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *J. Mem. Lang.* 40, 374–408. doi: 10.1006/jmla.1998.2618
- Wang, W., Li, X., Ning, N., and Zhang, J. X. (2012). The nature of the homophone density effect: an ERP study with Chinese spoken monosyllable homophones. *Neurosci. Lett.* 516, 67–71. doi: 10.1016/j.neulet.2012.03.059
- Wiener, S., and Turnbull, R. (2016). Constraints of tones, vowels and consonants on lexical selection in mandarin Chinese. *Lang. Speech* 59, 59–82. doi: 10.1177/0023830915578000
- Yao, Y., and Sharma, B. (2017). What is in the neighborhood of a tonal syllable? Evidence from auditory lexical decision in Mandarin Chinese. *Proc. Linguist. Soc. Am.* 2, 45–1. doi: 10.3765/plsa.v2i0.4090