



OPEN ACCESS

EDITED BY

Adam Linson,
The Open University, United Kingdom

REVIEWED BY

Zhiyuan Wang,
Roku, Inc., United States
Tongning Wu,
China Academy of Information and
Communications Technology, China

*CORRESPONDENCE

Chan Hee Kim
✉ chanypia@snu.ac.kr
Chun Kee Chung
✉ chungc@snu.ac.kr

RECEIVED 04 April 2025

REVISED 13 December 2025

ACCEPTED 22 December 2025

PUBLISHED 13 January 2026

CITATION

Kim CH, Seo J-E, Seol J and Chung CK (2026)
Figure–ground relationship of voices in
musical structure modulates reciprocal
frontotemporal connectivity.
Front. Neurosci. 19:1605800.
doi: 10.3389/fnins.2025.1605800

COPYRIGHT

© 2026 Kim, Seo, Seol and Chung. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Figure–ground relationship of voices in musical structure modulates reciprocal frontotemporal connectivity

Chan Hee Kim^{1,2*}, Jeong-Eun Seo³, Jaeho Seol² and Chun Kee Chung^{2,4*}

¹Department of Physiology and Neuroscience, Dental Research Institute, Seoul National University School of Dentistry, Seoul, Republic of Korea, ²Human Brain Function Laboratory, Neuroscience Research Institute, Seoul National University, Seoul, Republic of Korea, ³Department of Musicology, Seoul National University, Seoul, Republic of Korea, ⁴Neuroscience Research Institute, Seoul National University Medical Research Center, Seoul, Republic of Korea

When listening to polyphonic music, we often perceive a melody as the figure against the ground of accompanying sounds. However, with repeated exposure, this figure–ground relationship may naturally shift, allowing the melody to recede into the ground. In a previous study, we found the consistent pattern of frontotemporal connectivity for the “Twinkle, Twinkle, Little Star” (TTLS) melody in the headings of two *Variations (II and IV)* in Mozart’s 12 Variations, K. 265, indicating that the TTLS melody, but not the different lower voices, was the figure. However, the frontotemporal connectivity pattern may change in the same phrases repeating in the two variations. In the current study, we examined how frontotemporal connectivity changes in the repeated phrases. In the results, the frontotemporal connectivity pattern between the two variations changed in the final phrase after repeated passages. This suggests that the shift in the figure–ground relationship persists, with the TTLS melody becoming less prominent while the lower voices become relatively more prominent. Additionally, frontotemporal connectivity was strongly correlated with temporofrontal connectivity in the opposite direction. Finally, our data indicate that TTLS melody-based and sensory-based processes in response to a switched figure–ground relationship, are incorporated into the bidirectional connections between frontotemporal and temporofrontal connectivity. Our study highlights the brain’s ability to reconfigure figure–ground relationships in the processing of musical voices.

KEYWORDS

figure–ground perception, musical structure, musical voices, effective connectivity, inferior frontal gyrus, superior temporal gyrus

Introduction

Humans can identify a specific melody within homophonic and polyphonic music because it often appears in a higher pitch range, making it easily distinguishable (Fujioka et al., 2005; Trainor et al., 2014). This phenomenon aligns with the figure–ground concept in Gestalt psychology (Köhler, 1967; Wagemans et al., 2012). Similar to a visual stimulus (Supplementary Figure 1), the melody can serve as the “figure,” while other voices constitute the “ground.” Additionally, listeners can sometimes shift their attention during a phrase, perceiving the melody as the background while other voices being dominant

instead (Ragert et al., 2014; Deutch, 2019). However, even when the figure-ground relationship favors the figure as the more perceptually dominant voice, the ground can remain perceptible (Bigand et al., 2000), and vice versa.

Musical structure, including pitch, tonality, and harmony, is learned through experience, understanding of musical structure facilitates the recognition and anticipation of patterns in familiar pieces (Narmour, 2000; Tillmann et al., 2000). When a familiar melody appears in a musical piece, it is easily recognized the figure in the musical structure. However, repeated exposure to the melody may alter the figure-ground relationship between the upper and lower voices (Taher et al., 2016), and this change may eventually lead to the natural collapse of the figure-ground relationship centered on the upper voice of the familiar melody.

In our previous study using Mozart's 12 Variations, K. 265 (Kim et al., 2020), we observed that only frontotemporal connectivity between the left Heschl's gyrus (HG) and left inferior frontal gyrus (IFG) changed in response to the presence or absence of the "Twinkle, Twinkle, Little Star" (TTLS) melody of "C5-C5-G5-G5-A5." This connectivity pattern for the TTLS was observed across a target phrase (T) of 2.1 s at the beginning of each variation. However, if the figure-ground relationship shifts after repetitions, the connectivity strength can become inconsistent (Supplementary Figure 1C). The present study examined how frontotemporal connectivity for *Variations II* and *IV* (Figure 1) changes across four target phrases (T1-T4) featuring the TTLS melody. We hypothesized that: (1) If the connectivity pattern does not differ significantly between *Variations II* and *IV*, the TTLS melody remains the figure, with the lower voice serving as the ground; and (2) If the connectivity pattern differs significantly after repetitions, the TTLS melody may not be the sole figure, as the lower voices influence its prominence.

Materials and methods

Participants

In magnetoencephalography (MEG) recording, participants comprised 25 healthy individuals, all non-musicians, 15 women and 10 men with a mean age of 26.8 ± 3.4 years old. None had received formal musical training. All participants were right-handed, with a mean Edinburgh Handedness coefficient of 95.7 ± 7.1 . The study adhered to the principles of the Declaration of Helsinki and received approval from the Institutional Review Board of the Clinical Research Institute at Seoul National University Hospital (IRB No. C-1003-015-311). The research procedures adhered to relevant ethical guidelines and regulations. All participants provided informed, written consent after receiving a clear explanation of the study's purpose, procedures, potential risks, and benefits.

Stimuli

Mozart's K. 265 consists of the theme "Ah! Vous dirai-je Maman" and 12 variations (Supplementary Figure 2). *Variations I-IV* contain rhythmic, melodic, and textural variations on the theme.

Relative to the theme, the rhythmic patterns in the upper voices are transformed in *Variations I* and *III* and moved to the lower voices in *Variations II* and *IV*, sharing the TTLS melody (Figure 1A). This study focused on *Variations II* and *IV*, which share the TTLS melody but have differing lower parts, transforming through rhythmic changes to 8th note triplets and 16th notes (semiquavers), respectively. The tonality and harmonic structure remain the same for both variations. Each variation is based on the ternary form of $A (a + a) + B (b + a') + B (b + a')$. Phrases, including the "C5-C5-G5-G5-A5" melody, were repeated four times in each variation (Figure 1B). In this study, the term "variation" refers both to the musical form and to individual movements within that form, such as *Variation II* and *Variation IV*.

Recording

In a magnetically shielded room, the participants listened to Mozart's K. 265 while watching a silent movie clip (Love Actually, 2003, Universal Pictures, USA) for approximately 5 min. Musical stimuli were generated using STIM²™ (Neuroscan, Charlotte, NC, USA) and presented binaurally at 100 dB through MEG-compatible earphones (Tip-300, Nicolet, Madison, WI, USA). MEG signals were recorded using a 306-channel whole-head MEG system (Elekta Neuromag Vector View™, Helsinki, Finland) with a sampling frequency of 1,000 Hz and a bandpass filter of 0.1–200 Hz. The environmental magnetic noise in raw MEG signals was eliminated using the temporal signal space separation algorithm (Tesche et al., 1995; Taulu and Hari, 2009) implemented in MaxFilter 2.1.13 (Elekta Neuromag Oy, Helsinki, Finland). Electrooculogram, electrocardiogram, and muscle artifacts were also removed using independent component analysis. Participants did not perform behavioral tasks related to attentional shifts between voices during or after MEG recording. They also received no instructions regarding attentional focus during MEG recording.

Analysis

MEG source analysis

The MEG source signals of epochs from -100 to $2,100$ ms after the onset of each condition for four regional sources of bilateral HGs and IFGs, bandpass filtered at 14–30 Hz, were extracted using BESA 5.1.8.10 (MEGIS Software GmbH, Gräfelfing, Germany) after electrooculograms, electrocardiograms, and muscle artifacts were removed. Standard Talairach coordinates (x, y, and z in mm) for bilateral HGs (transverse, BA 41, BA 42) and IFGs (triangular part, BA 45) across participants were adapted from previous research (Kim et al., 2020). The coordinates were as follows: left HG (-53.5 , -30.5 , and 12.6), right HG (55.4 , -30.5 , and 12.6), left IFG (-55.5 , 11.7 , and 20.6), and right IFG (53.5 , 12.7 , and 20.6).

Time window

Variations II and *IV* were chosen as the two conditions for estimating connectivity differences, including the identical TTLS melody. Each Time window was 2,100 ms long, incorporating

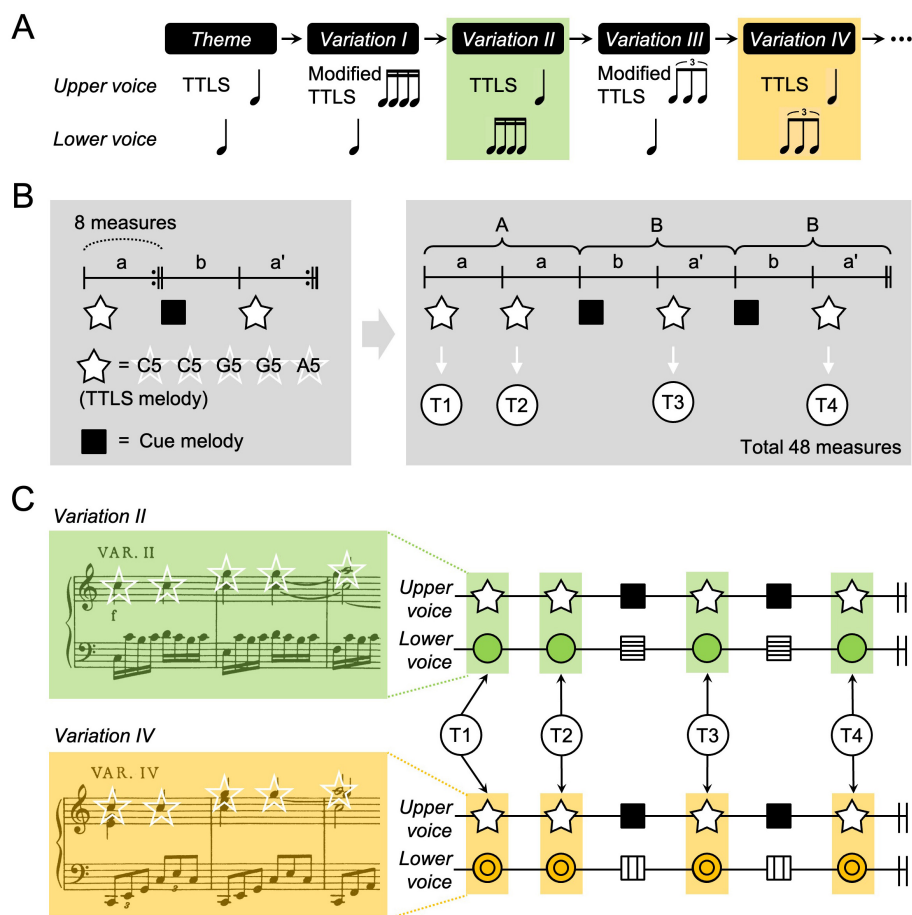


FIGURE 1

Musical stimuli. (A) In Mozart's 12 Variations, K. 265, the TTLS melody in the theme is modified in *Variations I* and *III* but not in *Variations II* or *IV*. The rhythmic appearance is the same within the pair of *Variations I* and *II* or that of *Variations III* and *IV*. (B) The left panel illustrates the structure of each variation involving repeat signs on the score. The right panel depicts the entire structure, comprising 48 measures of A (a + a) + B (b + a') + B (b + a'), as it is played. White and black squares denote TTLS and cue melodies, respectively, repeated four times per variation. (C) The target phrases (T1-T4) are highlighted using green and orange shaded boxes with the "C5-C5-G5-G5-A5" melody marked with white-lined stars. Both *Variations II* and *IV* have two streams of upper and lower voices. The lower voices in the target phrases show melodic and rhythmic variations on the theme, but the upper voice remains consistent. Details about the formation of the adapted and the full scores are shown in [Supplementary Figure 3](#). Musical scores of *Variation II* and *IV* were adapted from NMA Online: Neue Mozart-Ausgabe: Digitized Version (https://dme.mozarteum.at/DME/nma/nmapub_srch.php?l=2). TTLS, Twinkle Twinkle Little Star; T, target phrase.

the "C5-C5-G5-G5-A5" melody, as in a previous study (Kim et al., 2020). Time windows of 2,100 ms appeared four times per variation, labeled as T1, T2, T3, and T4 (Figure 1C and [Supplementary Figures 2, 3](#)). The lower voices accompanied by the "C5-C5-G5-G5-A5" melody differed between the two variations.

LTDMI analysis

Effective connectivity across target phrases between the two variations was measured using linearized time-delayed mutual information (LTDMI; Jin et al., 2010; Kim et al., 2021), a measure used in our previous study (Kim et al., 2020). LTDMI estimates the directionality of information transmission between the time series of two regional sources, enabling the observation of interhemispheric and interregional connectivity, which is essential for processing musical elements in bilateral IFGs and HGs. While our primary focus was the regional connection from the left IFG

to the right HG, we verified our results for all connections among the bilateral IFGs and HGs. The effective connectivity for the 12 connections between regional sources of the bilateral HGs and IFGs was estimated using MATLAB 7.7.0.471 (Math Works Inc., Natick, MA, USA; see also [Supplementary Table 1](#) for individual LTDMI values calculated for 12 connections). For each subject, the mean LTDMI for the 2,100-ms epoch was calculated for each of 4 target phrases (T1, T2, T3, and T4) × 2 variations (*Variations II* and *IV*).

Statistics

Statistical comparisons of mean LTDMI values for *Variations II* and *IV* were performed using SPSS 21.0 software (IBM, Armonk, NY, USA). For the mean LTDMI values in four target phrases, we conducted the non-parametric Wilcoxon signed-rank test due to the non-Gaussian distribution of LTDMI data. In each

case, the significance level (α) for rejecting the null hypothesis (H_0 , indicating no difference between *Variation II* and *Variation IV* in the mean LTDMI values), was 0.05. In addition, in the nonparametric Spearman correlation test for each pair between the frontotemporal connectivity difference value [Left IFG \rightarrow Right HG(*Variation IV*–*Variation II*)] and the other 11 connectivity difference values, except of Left IFG \rightarrow Right HG among 12 connections between the bilateral IFGs and HGs, the Type I errors that were caused by multiple comparisons among the 11 connection pairs in the Spearman correlation test were adjusted by the Bonferroni test.

Results

LTDMI differences between two variations for four target phrases

The frontotemporal connectivity from the left IFG to the right HG between *Variations II* and *IV* was calculated for four target phrases (T1–T4), each repeated four times per variation (Figure 1). We independently performed a Wilcoxon signed-rank test for the LTDMI values of each target phrase to confirm the changes in the frontotemporal connectivity between the two variations. The difference between the two variations was significant only in T4 among four target phrases, as indicated by the Wilcoxon signed-rank test ($Z = -2.112$, $P = 0.035$; Figure 2A). In T4, frontotemporal connectivity was enhanced in *Variation IV* compared with *Variation II*. However, significant differences were not observed in T1–T3 ($P > 0.05$ in all cases). In addition, we confirmed that, among 12 connections between the bilateral IFGs and HGs, the only significant result corresponded specially to frontotemporal connectivity from the left IFG to the right HG (Supplementary Table 1). We observed a near-significant effect in temporofrontal connectivity from the right HG to the left IFG, in the opposite direction of frontotemporal connectivity (Right HG \rightarrow Left IFG, $Z = -1.843$, $P = 0.065$; Figure 2A and Supplementary Table 1), which was not initially predicted in our hypothesis. The significance level (α) for the null hypothesis was independently tested for each target phrase (T1–T4), since the target phrases of T1–T4 existed in completely different musical contexts within the formal structure of ternary form. Additionally, comparisons between target phrases within a variation were not considered as a hypothesis.

Correlation between frontotemporal and temporofrontal connectivity

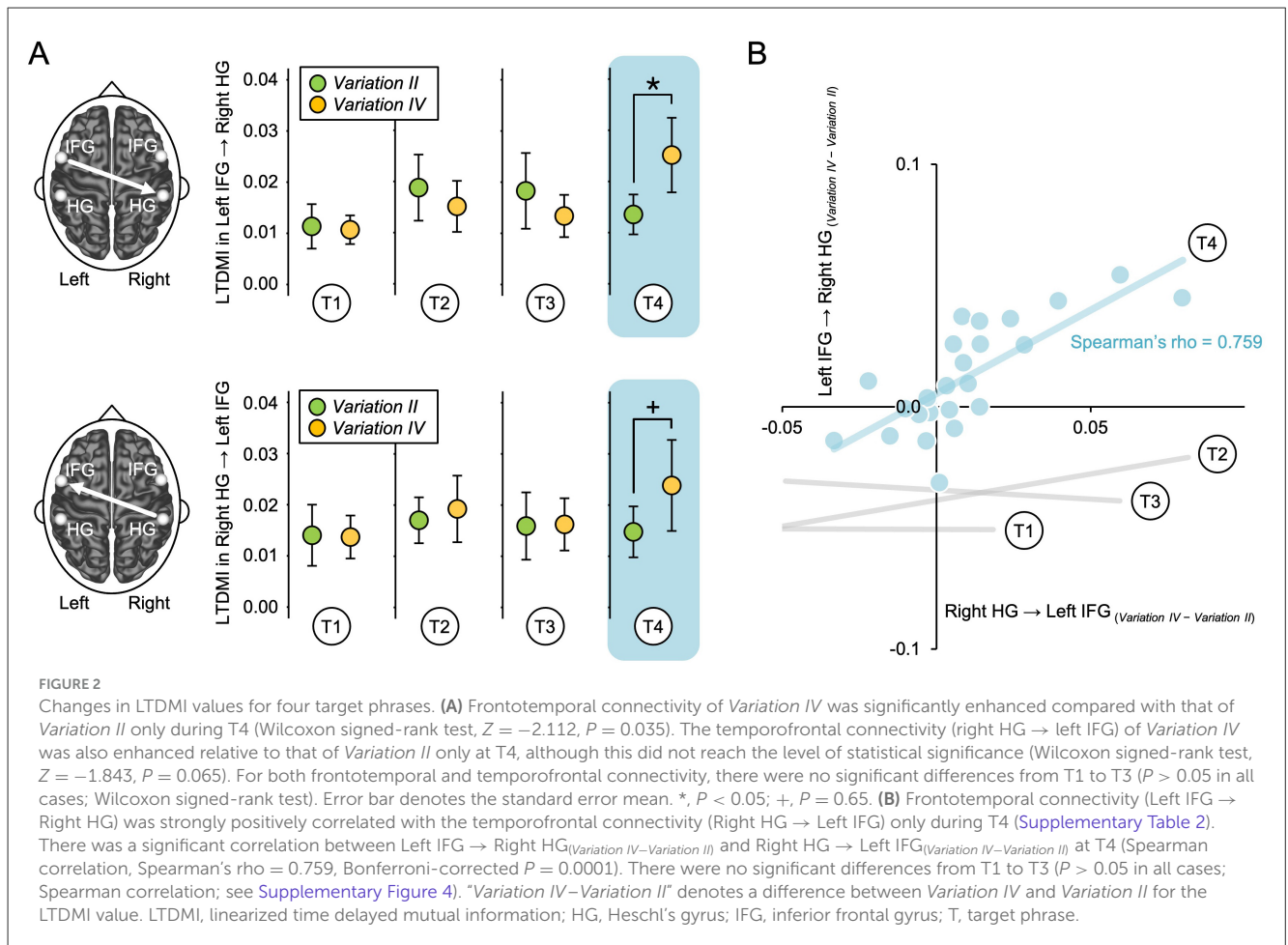
Correlation analyses were conducted to confirm (1) whether a similar pattern between frontotemporal and temporofrontal connectivity refers to bidirectional information transmission between the left IFG and the right HG and (2) whether a similar pattern is only specialized in the temporofrontal connectivity (Right HG \rightarrow Left IFG) among 12 connections between the bilateral IFGs and HGs, which are key areas for the music process. To perform this estimation, we first computed the difference values between *Variations II* and *Variation IV* for the LTDMI values in 12 connections between the bilateral IFGs

and HGs for all target phrases of T1–T4. Next, we estimated the correlation between the frontotemporal connectivity difference value [Left IFG \rightarrow Right HG(*Variation IV*–*Variation II*)], with 11 other connectivity difference values. In the Spearman correlation test result, a significant correlation was only observed between Left IFG \rightarrow Right HG(*Variation IV*–*Variation II*) and Right HG \rightarrow Left IFG(*Variation IV*–*Variation II*) for T4, among 44 combinations of 11 connections \times 4 target phrases (Spearman's $\rho = 0.759$, Bonferroni-corrected $P = 0.0001$; Figure 2B, Supplementary Figure 4, and Supplementary Table 2). The frontotemporal connectivity (Left IFG \rightarrow Right HG) was strongly positively correlated with the temporofrontal connectivity (Right HG \rightarrow Left IFG), reflecting similar information processing in *Variation II* and *IV*.

Discussion

A difference in frontotemporal connectivity from the left IFG to the right HG between the two variations was only observed in the final target phrase of T4 and not in the preceding three phrases (Figure 2A). As we hypothesized, frontotemporal connectivity showed inconsistency in the figure–ground relationship between the two variations of the TTLS melody in a repeated phrase of T4. This indicates that the perceptual dominance of the TTLS melody in voice perception was weakened. Each variation included cue phrases such as “Up above the world...,” predicting the recurrence of the TTLS melody (Supplementary Figure 2). The two musical structures—variation and ternary forms—establish global and regional contexts, respectively. In each variation, T4 is introduced after the second iteration of the cue phrase within the regional ternary context, facilitating anticipation of melodic recurrence. Moreover, at the global level of the variation form, the same structure involving T4 is repeated in *Variation II* and *Variation IV*, further enhancing anticipatory processing. The training of the repeated upper voice and its perceptual prominence might have facilitated participants' recognition of the lower voice (Taher et al., 2016). The LTDMI value was higher in *Variation IV* than in *Variation II* during T4. We interpret that the connectivity reduction in *Variation II* for T4 is attributable to the properties of its lower voice, which differed from them in *Variation IV* (Figure 3). Our findings show that participants did not solely focus on the TTLS melody at T4 but could also detect other sounds.

In our previous study (Kim et al., 2020), unidirectional transmission of information from the left IFG to the right HG, showing frontotemporal connectivity, was associated with the recognition of the TTLS melody in the heading of each variation. However, the involvement of temporofrontal connectivity, which was in the opposite direction to the frontotemporal connectivity, was observed in repeated phrases in musical context evoked (Figure 2B). The temporofrontal connectivity was strongly correlated with the frontotemporal connectivity in T4 (Figure 2B and Supplementary Table 2). The top–down processing of a familiar melody can significantly influence the figure–ground relationship (Strüber and Stadler, 1999; Nelson and Palmer, 2007). Considering that the roles of frontotemporal and temporofrontal connectivity are linked to the TTLS melody-based and sensory-based processes, respectively, this heightened



connectivity could indicate a dual process: extracting the novel lower voice in the target phrase, including the familiar TTLS melody, and dissecting the components within the novel lower voices. The temporofrontal connectivity supports the frontotemporal connectivity. Thus, bidirectional connectivity of the frontotemporal and temporofrontal pathways between the left IFG and the right HG is possibly modulated by both a top-down process based on knowledge of the TTLS melody and a bottom-up process based on new information on the voices accumulated while sequentially listening to target and cue phrases in each variation (Alho et al., 2015; Dzafic et al., 2021).

Numerous researches on the auditory figure-ground relationship have conducted auditory scene analysis (Bregman, 1994) and grouping (Wagemans et al., 2012) using tasks that discriminate a sound pattern as a figure from a ground of tone and chord sequences with irregularities in spectral and temporal properties (Teki et al., 2011; O'Sullivan et al., 2015; Toth et al., 2016). Functional magnetic resonance imaging and electroencephalography studies (Teki et al., 2011; O'Sullivan et al., 2015; Toth et al., 2016) have reported that regions involved in discriminating this figure-ground perception comprise the primary auditory area, superior temporal sulcus, superior temporal gyrus, intraparietal sulcus, medial/superior frontal gyrus, and cingulate cortex. The processing of multiple voices in this study involved the IFG and the HG, both exhibiting information transmission.

The left IFG is crucial in processing familiarity (Plailly et al., 2007) and in music-syntactic processing, indicating implicit learning (Sammler et al., 2011) and contributing to memory retrieval (Platel et al., 2003; Watanabe et al., 2008). Moreover, the left IFG is involved in differentiating between melody and accompaniment (Spada et al., 2014). The left IFG is highly activated during the recognition of pattern deviations in melody (Habermeyer et al., 2009), musical novelty in a context (Tillmann et al., 2003), and conscious experience (Weilhammer et al., 2021). The left IFG is also associated with semantic and syntactic processing (Zhu et al., 2022), and sentence comprehension (van der Burght et al., 2019). In contrast, the right HG predominantly processes tone deviance (Sabri et al., 2006; Nan and Friederici, 2013) and the segregation of auditory streams (Snyder et al., 2006). The right auditory cortex is the dominant site for music processing (Perani et al., 2010), auditory stream segregation (Snyder et al., 2006), and spectral pitch (Schneider et al., 2005).

The IFG and HG are pivotal areas for music perception, and their connectivity is discussed in relation to syntax processes (Papoutsis et al., 2011; Kim et al., 2019, 2021), categorization (Roswadowski et al., 2021), and the working memory of the melody process (Burunat et al., 2014). The temporofrontal network is engaged in the categorization of vocal signals (Roswadowski et al., 2021), while frontotemporal connection is involved in both top-down and bottom-up processes in sensory learning (Dzafic

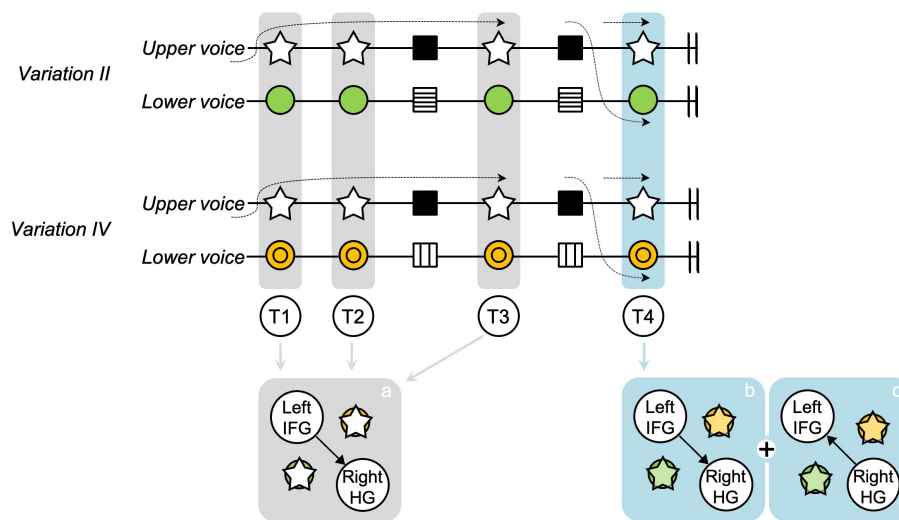


FIGURE 3

Figure-ground relationship between voices according to frontotemporal and temporofrontal connectivity changes. The figure shows how connectivity between the left IFG and the right HG changes from T1 to T4 when the same passages are repeated in each variation. Frontotemporal connectivity from the left IFG to the right HG, the TTLS connectivity, is consistent from T1 to T3, focusing on the figure of the upper voice. However, at T4, a different frontotemporal connectivity pattern is exhibited between *Variations II* and *IV*. This indicates that frontotemporal connectivity may no longer be related to the TTLS melody but rather to the lower voice, reflecting a shift away from TTLS-specific connectivity. Lower voices, masked by the perceptual dominance of the TTLS melody until T3, may have become audible alongside the TTLS melody at T4. However, because these results do not demonstrate whether figure and ground are perceptually separated or integrated, predictable bidirectional processes are represented using overlapping circles and stars. HG, Heschl's gyrus; IFG, inferior frontal gyrus; T, target phrase; TTLS, Twinkle Twinkle Little Star.

et al., 2021). Directional information flows within frontotemporal connectivity may therefore explain how the left IFG and the right HG collaborate in processing target phrases. The involvement of the left IFG and right HG may reflect the entire process of naturally grasping, comparing, and understanding the voices in target phrases within a particular context rather than simply perceiving them as sounds. Accordingly, enhanced frontotemporal and temporofrontal connectivity may reflect the integrated processes by which the brain recognizes the TTLS melody based on memory, segregates the melody, and detects differences in the lower voices relative to the prior context.

Previous studies have selectively manipulated stimuli or directed participants' attention to specific auditory streams (Uhlig et al., 2013; Ragert et al., 2014; Spada et al., 2014; Strait et al., 2015; Hausfeld et al., 2018; Puschmann et al., 2019; Barrett et al., 2021). Attention has been shown to be critical for figure-ground perception (Poort et al., 2012). Therefore, research on such perception uses artificially composed stimuli to direct participants' attention. During the MEG experiment in this study, all participants passively listened to the naturalistic music of Mozart's 12 Variations, K. 265, without any instructions regarding focusing their attention on a specific voice or melody. Thus, whether both figure and ground were processed attentively or pre-attentively remains unclear, given the absence of intentional attention control. Although listeners may focus on a particular voice while listening to music, the changing flow of music can encompass brief perceptible moments in which the figure-ground relationship continuously shifts without listeners consciously realizing it. Indeed, music listeners can automatically process information such as syntactic errors and tone deviations without intentional attention (Maess et al., 2001; Naatanen et al., 2007). We interpreted that participants could attentively or pre-attentively detect sonic changes in the

voices at that moment, although our data do not confirm that non-musicians could perceptually segregate the streams or identify which voice evoked the sonic differences (Figure 3). Our results successfully captured the moment when the figure-ground relationship between the upper and lower voices changed, as evidenced by the difference in frontotemporal connectivity for repeated phrases in the two variations and the correlation between frontotemporal and temporofrontal connectivity.

Familiarity is critical for explaining the figure-ground experiment (Palmer, 1999; Hulleman and Humphreys, 2004; Nelson and Palmer, 2007). The target phrases in our stimuli involved the TTLS song, which has been used in studies related to the perception of familiar melodies (Trehub et al., 1985; Uptis, 1990; Besson et al., 1994; Creel, 2019). Familiarity would be naturally implied in the theme and all of its variations, considering that Mozart's 12 Variations, K. 265, is based on the TTLS melody. Logically, the familiarity implied in the TTLS melody might influence participants' figure-ground perception. However, in our previous study focusing on the TTLS melody (Kim et al., 2020), we could not directly prove the effect of familiarity on frontotemporal connectivity as the connectivity changed irrespective of the presence or absence of the TTLS melody. In our present study, the same TTLS melody appeared repeatedly in *Variations II* and *IV*. The effects of familiarity are consistent in both *Variations II* and *IV*. Thus, the TTLS melody was used to assess changes in the figure-ground relationship.

Naturalistic stimuli have been used to examine various topics (Saarimäki, 2021; Izen et al., 2023; Tervaniemi, 2023). In studies on the concepts of emotion (Singer et al., 2016; Putkinen et al., 2021), melodic expectation (Kern et al., 2022), temporal aspects of rhythm and beat (Sturm et al., 2015; Weineck et al., 2022), and familiarity (Leaver et al., 2009), multiple naturalistic pieces

have been used as musical stimuli. Some studies using naturalistic stimuli have examined their hypotheses on topics such as motif, musical features, timbre, and depression, based on a single piece (Alluri et al., 2012; Cong et al., 2013; Burunat et al., 2016; Liu et al., 2020). Our hypothesis was also created for the melody of TTLS and the figure-ground relationship of voices using Mozart's 12 Variations, K. 265. In the fields of audition (Uhlir et al., 2013; Ragert et al., 2014; Hausfeld et al., 2018; Barrett et al., 2021) and vision (Peterson et al., 1991; Lamme, 1995; Super et al., 2003; Zhang and Von Der Heydt, 2010; Von der Heydt, 2015), changes in the figure-ground relationship between the elements of an object can be simply observed and explained by comparing the related objects. However, naturalistic music has its own narrative, which can be described by its structure. The connectivity reflects complicated processes for the target phrases of 2.1 s without the context and for each 2.1-s-long target phrases in the theme and variations, leading up to the target phrase. This approach, however, may constitute a critical shortcoming of our study, compared with measurements using artificially composed stimuli. Furthermore, no verbal reports or other measures were obtained during or after MEG recording to determine what participants perceived as the "figure" in the music at each moment, including critical time windows. As a result, connectivity changes were measured under naturalistic listening conditions in which participants passively listened to music. These limitations can be addressed in future studies using novel paradigms that incorporate detailed behavioral responses and larger sample sizes. Nevertheless, our results reflected the humans' ubiquitous experiences with individual participants.

In addition to the use of naturalistic music, our study had other limitations. While non-musicians can perceive the figure separately from the ground (Toth et al., 2016), they tend to focus more on the upper voice than on a lower one (Sloboda and Edworthy, 1981). In contrast, musicians are more sensitive to voice perception and are better able to distinguish voices (Fujioka et al., 2005; Strait et al., 2015). Music training significantly influences selective attention (Puschmann et al., 2019). We did not recruit musicians as participants to examine our hypothesis in terms of basic musical ability. Recruiting non-musicians might have impacted our results. The temporofrontal connectivity may show a statistically significant distinction ($P < 0.05$) since the recognition of different elements in the lower voices (via temporofrontal connectivity) involves a more complex cognitive process. Therefore, future studies should verify these results with musicians. Furthermore, this study concentrated on frontal and temporal regions. Our findings should be verified at the whole-brain level. In the current experimental paradigm, we did not consider an additional test individual preference of subjects. The preference is also influential of music listening, and which should be addressed importantly in further studies with a novel experimental paradigm. The MEG recording and analysis approach used in this study may be replicable using other modalities, such as EEG. Despite these limitations, the use of Mozart's 12 Variations, K. 265, was invaluable in understanding the fundamental neural processes associated with processing real music featuring multiple voices and elucidating the human experience of music. Our findings elucidated how the brain dissects voices from the multidimensional structures of music and reconstructs the figure-ground relationship between voices.

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 Institutional Review Board of the Clinical Research Institute at Seoul National University Hospital (IRB No. C-1003-015-311). 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

CK: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. J-ES: Methodology, Writing – review & editing. JS: Investigation, Writing – review & editing, Methodology. CC: Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This research was supported by Samsung Research Funding & Incubation Center for Future Technology (SRFC-IT1902-08, Decoding Inner Music Using Electroencephalography), and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science & ICT (NRF-2021R1A4A200180312) and the Ministry of Education (RS-2022-NR075566).

Acknowledgments

We sincerely appreciate Ji Hyang Nam for her technical support in MEG data acquisition.

Conflict of interest

The author declares 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) declared that generative AI was used in the creation of this manuscript. Furthermore, I acknowledge the use of ChatGPT (OpenAI Version 2) for assistance with language editing during the preparation for the original draft 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/fnins.2025.1605800/full#supplementary-material>

References

- Alho, K., Salmi, J., Koistinen, S., Salonen, O., and Rinne, T. (2015). Top-down controlled and bottom-up triggered orienting of auditory attention to pitch activate overlapping brain networks. *Brain Res.* 1626, 136–145. doi: 10.1016/j.brainres.2014.12.050
- Alluri, V., Toiviainen, P., Jaaskelainen, I. P., Gleason, E., Sams, M., and Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage* 59, 3677–3689. doi: 10.1016/j.neuroimage.2011.11.019
- Barrett, K. C., Ashley, R., Strait, D. L., Skoe, E., Limb, C. J., and Kraus, N. (2021). Multi-voiced music bypasses attentional limitations in the brain. *Front. Neurosci.* 15:588914. doi: 10.3389/fnins.2021.588914
- Besson, M., Faïta, F., and Requin, J. (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neurosci. Lett.* 168, 101–105. doi: 10.1016/0304-3940(94)90426-X
- Bigand, E., McAdams, S., and Forêt, S. (2000). Divided attention in music. *Int. J. Psychol.* 35, 270–278. doi: 10.1080/002075900750047987
- Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press. doi: 10.1121/1.408434
- Burunat, I., Alluri, V., Toiviainen, P., Numminen, J., and Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex* 57, 254–269. doi: 10.1016/j.cortex.2014.04.012
- Burunat, I., Toiviainen, P., Alluri, V., Bogert, B., Ristaniemi, T., Sams, M., et al. (2016). The reliability of continuous brain responses during naturalistic listening to music. *Neuroimage* 124, 224–231. doi: 10.1016/j.neuroimage.2015.09.005
- Cong, F., Alluri, V., Nandi, A. K., Toiviainen, P., Fa, R., Abu-Jamous, B., et al. (2013). Linking brain responses to naturalistic music through analysis of ongoing EEG and stimulus features. *IEEE Trans. Multimedia* 15, 1060–1069. doi: 10.1109/TMM.2013.2253452
- Creel, S. C. (2019). The familiar-melody advantage in auditory perceptual development: Parallels between spoken language acquisition and general auditory perception. *Attention Percept. Psychophys.* 81, 948–957. doi: 10.3758/s13414-018-01663-7
- Deutch, D. (2019). *Chapter 3. The Perceptual Organization of Streams of Sound*. Oxford: Oxford University Press. doi: 10.1093/oso/9780190206833.003.0004
- Dzafic, I., Larsen, K. M., Darke, H., Pertile, H., Carter, O., Sundram, S., et al. (2021). Stronger top-down and weaker bottom-up frontotemporal connections during sensory learning are associated with severity of psychotic phenomena. *Schizophr. Bull.* 47, 1039–1047. doi: 10.1093/schbul/sbaa188
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., and Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *J. Cogn. Neurosci.* 17, 1578–1592. doi: 10.1162/089892905774597263
- Habermeyer, B., Herdener, M., Esposito, F., Hilti, C. C., Klarhofer, M., di Salle, F., et al. (2009). Neural correlates of pre-attentive processing of pattern deviance in professional musicians. *Hum. Brain Mapp.* 30, 3736–3747. doi: 10.1002/hbm.20802
- Hausfeld, L., Riecke, L., Valente, G., and Formisano, E. (2018). Cortical tracking of multiple streams outside the focus of attention in naturalistic auditory scenes. *Neuroimage* 181, 617–626. doi: 10.1016/j.neuroimage.2018.07.052
- Hulleman, J., and Humphreys, G. W. (2004). A new cue to figure-ground coding: top-bottom polarity. *Vision Res.* 44, 2779–2791. doi: 10.1016/j.visres.2004.06.012
- Izen, S. C., Cassano-Coleman, R. Y., and Piazza, E. A. (2023). Music as a window into real-world communication. *Front. Psychol.* 14:1012839. doi: 10.3389/fpsyg.2023.1012839
- Jin, S. H., Lin, P., and Hallett, M. (2010). Linear and nonlinear information flow based on time-delayed mutual information method and its application to corticomuscular interaction. *Clin. Neurophysiol.* 121, 392–401. doi: 10.1016/j.clinph.2009.09.033
- Kern, P., Heilbron, M., de Lange, F. P., and Spaak, E. (2022). Cortical activity during naturalistic music listening reflects short-range predictions based on long-term experience. *Elife* 11:e80935. doi: 10.7554/eLife.80935.sa2
- Kim, C. H., Jin, S. H., Kim, J. S., Kim, Y., Yi, S. W., and Chung, C. K. (2021). Dissociation of connectivity for syntactic irregularity and perceptual ambiguity in musical chord stimuli. *Front. Neurosci.* 15:693629. doi: 10.3389/fnins.2021.693629
- Kim, C. H., Kim, J. S., Choi, Y., Kyong, J. S., Kim, Y., Yi, S. W., et al. (2019). Change in left inferior frontal connectivity with less unexpected harmonic cadence by musical expertise. *PLoS ONE* 14:e0223283. doi: 10.1371/journal.pone.0223283
- Kim, C. H., Seol, J., Jin, S.-H., Kim, J. S., Kim, Y., Yi, S. W., et al. (2020). Increased fronto-temporal connectivity by modified melody in real music. *PLoS ONE* 15:e0235770. doi: 10.1371/journal.pone.0235770
- Köhler, W. (1967). Gestalt psychology. *Psychologische Forschung* 31, XVIII–XXX. doi: 10.1007/BF00422382
- Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* 15, 1605–1615. doi: 10.1523/JNEUROSCI.15-02-01605.1995
- Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., and Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *J. Neurosci.* 29, 2477–2485. doi: 10.1523/JNEUROSCI.4921-08.2009
- Liu, W., Zhang, C., Wang, X., Xu, J., Chang, Y., Ristaniemi, T., et al. (2020). Functional connectivity of major depression disorder using ongoing EEG during music perception. *Clin. Neurophysiol.* 131, 2413–2422. doi: 10.1016/j.clinph.2020.06.031
- Maess, B., Koelsch, S., Gunter, T. C., and Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.* 4, 540–545. doi: 10.1038/87502
- Naatanen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026
- Nan, Y., and Friederici, A. D. (2013). Differential roles of right temporal cortex and Broca's area in pitch processing: evidence from music and Mandarin. *Hum. Brain Mapp.* 34, 2045–2054. doi: 10.1002/hbm.22046
- Narmour, E. (2000). Music expectation by cognitive rule-mapping. *Music Percept.* 17, 329–398. doi: 10.2307/40285821
- Nelson, R. A., and Palmer, S. E. (2007). Familiar shapes attract attention in figure-ground displays. *Percept. Psychophys.* 69, 382–392. doi: 10.3758/BF03193759
- O'Sullivan, J. A., Shamma, S. A., and Lalor, E. C. (2015). Evidence for neural computations of temporal coherence in an auditory scene and their enhancement during active listening. *J. Neurosci.* 35, 7256–7263. doi: 10.1523/JNEUROSCI.4973-14.2015
- Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.
- Papoutsis, M., Stamatakis, E. A., Griffiths, J., Marslen-Wilson, W. D., and Tyler, L. K. (2011). Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. *Neuroimage* 58, 656–664. doi: 10.1016/j.neuroimage.2011.06.036
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., et al. (2010). Functional specializations for music processing in the human newborn brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4758–4763. doi: 10.1073/pnas.0909074107
- Peterson, M. A., Harvey, E. M., and Weidenbacher, H. J. (1991). Shape recognition contributions to figure-ground reversal: which route counts? *J. Exp. Psychol. Hum. Percept. Perform.* 17:1075. doi: 10.1037//0096-1523.17.4.1075
- Plailly, J., Tillmann, B., and Royet, J. P. (2007). The feeling of familiarity of music and odors: the same neural signature? *Cereb. Cortex* 17, 2650–2658. doi: 10.1093/cercor/bhl173

- Platel, H., Baron, J. C., Desgranges, B., Bernard, F., and Eustache, F. (2003). Semantic and episodic memory of music are subserved by distinct neural networks. *Neuroimage* 20, 244–256. doi: 10.1016/S1053-8119(03)00287-8
- Poort, J., Raudies, F., Wannig, A., Lamme, V. A., Neumann, H., and Roelfsema, P. R. (2012). The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* 75, 143–156. doi: 10.1016/j.neuron.2012.04.032
- Puschmann, S., Baillet, S., and Zatorre, R. J. (2019). Musicians at the Cocktail party: neural substrates of musical training during selective listening in multispeaker situations. *Cereb. Cortex* 29, 3253–3265. doi: 10.1093/cercor/bhy193
- Putkinen, V., Nazari-Farsani, S., Seppälä, K., Karjalainen, T., Sun, L., Karlsson, H. K., et al. (2021). Decoding music-evoked emotions in the auditory and motor cortex. *Cereb. Cortex* 31, 2549–2560. doi: 10.1093/cercor/bhaa373
- Ragert, M., Fairhurst, M. T., and Keller, P. E. (2014). Segregation and integration of auditory streams when listening to multi-part music. *PLoS ONE* 9:e84085. doi: 10.1371/journal.pone.0084085
- Roswandowitz, C., Swanborough, H., and Fruhholz, S. (2021). Categorizing human vocal signals depends on an integrated auditory-frontal cortical network. *Hum. Brain Mapp.* 42, 1503–1517. doi: 10.1002/hbm.25309
- Saarimäki, H. (2021). Naturalistic stimuli in affective neuroimaging: a review. *Front. Hum. Neurosci.* 15:675068. doi: 10.3389/fnhum.2021.675068
- Sabri, M., Liebenthal, E., Waldron, E. J., Medler, D. A., and Binder, J. R. (2006). Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI study. *J. Cogn. Neurosci.* 18, 689–700. doi: 10.1162/jocn.2006.18.5.689
- Sammler, D., Koelsch, S., and Friederici, A. D. (2011). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex* 47, 659–673. doi: 10.1016/j.cortex.2010.04.007
- Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., et al. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.* 8, 1241–1247. doi: 10.1038/nn1530
- Singer, N., Jacoby, N., Lin, T., Raz, G., Shpigelman, L., Gilam, G., et al. (2016). Common modulation of limbic network activation underlies musical emotions as they unfold. *Neuroimage* 141, 517–529. doi: 10.1016/j.neuroimage.2016.07.002
- Sloboda, J., and Edworthy, J. (1981). Attending to two melodies at once: the of key relatedness. *Psychol. Music* 9, 39–43. doi: 10.1177/03057356810090010701
- 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
- Spada, D., Verga, L., Iadanza, A., Tettamanti, M., and Perani, D. (2014). The auditory scene: an fMRI study on melody and accompaniment in professional pianists. *Neuroimage* 102(Pt 2), 764–775. doi: 10.1016/j.neuroimage.2014.08.036
- Strait, D. L., Slater, J., O'Connell, S., and Kraus, N. (2015). Music training relates to the development of neural mechanisms of selective auditory attention. *Dev. Cogn. Neurosci.* 12, 94–104. doi: 10.1016/j.dcn.2015.01.001
- Strüber, D., and Stadler, M. (1999). Differences in top-down influences on the reversal rate of different categories of reversible figures. *Perception* 28, 1185–1196. doi: 10.1068/p2973
- Sturm, I., Dahne, S., Blankertz, B., and Curio, G. (2015). Multi-variate EEG analysis as a novel tool to examine brain responses to naturalistic music stimuli. *PLoS ONE* 10:e0141281. doi: 10.1371/journal.pone.0141281
- Super, H., van der Togt, C., Spekreijse, H., and Lamme, V. A. (2003). Internal state of monkey primary visual cortex (V1) predicts figure-ground perception. *J. Neurosci.* 23, 3407–3414. doi: 10.1523/JNEUROSCI.23-08-03407.2003
- Taher, C., Rusch, R., and McAdams, S. (2016). Effects of repetition on attention in two-part counterpoint. *Music Percept. Interdiscipl. J.* 33, 306–318. doi: 10.1525/mp.2016.33.3.306
- Taulu, S., and Hari, R. (2009). Removal of magnetoencephalographic artifacts with temporal signal-space separation: demonstration with single-trial auditory-evoked responses. *Hum. Brain Mapp.* 30, 1524–1534. doi: 10.1002/hbm.20627
- Teeki, S., Chait, M., Kumar, S., von Kriegstein, K., and Griffiths, T. D. (2011). Brain bases for auditory stimulus-driven figure-ground segregation. *J. Neurosci.* 31, 164–171. doi: 10.1523/JNEUROSCI.3788-10.2011
- Tervaniemi, M. (2023). The neuroscience of music—towards ecological validity. *Trends Neurosci.* 46, 355–364. doi: 10.1016/j.tins.2023.03.001
- Tesche, C. D., Uusitalo, M. A., Ilmoniemi, R. J., Huotilainen, M., Kajola, M., and Salonen, O. (1995). Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. *Electroencephalogr. Clin. Neurophysiol.* 95, 189–200. doi: 10.1016/0013-4694(95)00064-6
- Tillmann, B., Bharucha, J. J., and Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychol. Rev.* 107:885. doi: 10.1037/0033-295X.107.4.885
- Tillmann, B., Janata, P., and Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Brain Res. Cogn. Brain Res.* 16, 145–161. doi: 10.1016/S0926-6410(02)00245-8
- Toth, B., Kocsis, Z., Haden, G. P., Szerafin, A., Shinn-Cunningham, B. G., and Winkler, I. (2016). EEG signatures accompanying auditory figure-ground segregation. *Neuroimage* 141, 108–119. doi: 10.1016/j.neuroimage.2016.07.028
- Trainor, L. J., Marie, C., Bruce, I. C., and Bidelman, G. M. (2014). Explaining the high voice superiority effect in polyphonic music: evidence from cortical evoked potentials and peripheral auditory models. *Hear. Res.* 308, 60–70. doi: 10.1016/j.heares.2013.07.014
- Trehub, S. E., Morrongiello, B. A., and Thorpe, L. A. (1985). Children's perception of familiar melodies: the role of intervals, contour, and key. *Psychomusicol. J. Res. Music Cogn.* 5:39. doi: 10.1037/h0094201
- Uhlig, M., Fairhurst, M. T., and Keller, P. E. (2013). The importance of integration and top-down salience when listening to complex multi-part musical stimuli. *Neuroimage* 77, 52–61. doi: 10.1016/j.neuroimage.2013.03.051
- Uptis, R. (1990). Children's invented notations of familiar and unfamiliar melodies. *Psychomusicol. J. Res. Music Cogn.* 9:89. doi: 10.1037/h0094156
- van der Burgh, C. L., Goucha, T., Friederici, A. D., Kreitewolf, J., and Hartwigsen, G. (2019). Intonation guides sentence processing in the left inferior frontal gyrus. *Cortex* 117, 122–134. doi: 10.1016/j.cortex.2019.02.011
- Von der Heydt, R. (2015). Figure-ground organization and the emergence of proto-objects in the visual cortex. *Front. Psychol.* 6:1695. doi: 10.3389/fpsyg.2015.01695
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., et al. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychol. Bull.* 138, 1172–1217. doi: 10.1037/a0029333
- Watanabe, T., Yagishita, S., and Kikyo, H. (2008). Memory of music: roles of right hippocampus and left inferior frontal gyrus. *Neuroimage* 39, 483–491. doi: 10.1016/j.neuroimage.2007.08.024
- Weinhammer, V., Fritsch, M., Chikermane, M., Eckert, A. L., Kanthak, K., Stuke, H., et al. (2021). An active role of inferior frontal cortex in conscious experience. *Curr. Biol.* 31, 2868–2880.e2868. doi: 10.1016/j.cub.2021.04.043
- Weineck, K., Wen, O. X., and Henry, M. J. (2022). Neural synchronization is strongest to the spectral flux of slow music and depends on familiarity and beat salience. *Elife* 11:e75515. doi: 10.7554/eLife.75515.sa2
- Zhang, N. R., and Von Der Heydt, R. (2010). Analysis of the context integration mechanisms underlying figure-ground organization in the visual cortex. *J. Neurosci.* 30, 6482–6496. doi: 10.1523/JNEUROSCI.5168-09.2010
- Zhu, Y., Xu, M., Lu, J., Hu, J., Kwok, V. P. Y., Zhou, Y., et al. (2022). Distinct spatiotemporal patterns of syntactic and semantic processing in human inferior frontal gyrus. *Nat. Hum. Behav.* 6, 1104–1111. doi: 10.1038/s41562-022-01334-6