Lateralization in the dichotic listening of tones is influenced by the content of speech

Ning Mei [a],[1], Adeen Flinker [b],[1], Miaomiao Zhu [c], Qing Cai [c],[d], Xing Tian [d],[e],*

[a] Basque Center on Cognition, Brain, and Language, Spain
[b] School of Medicine, New York University, United States
[c] Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), School of Psychology and Cognitive Science, East China Normal University, China
[d] NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, China
[e] Division of Arts and Sciences, New York University Shanghai, China

A B S T R A C T

Cognitive functions, for example speech processing, are distributed asymmetrically in the two hemispheres that mostly have homologous anatomical structures. Dichotic listening is a well-established paradigm to investigate hemispherical lateralization of speech. However, the mixed results of dichotic listening, especially when using tonal languages as stimuli, complicates the investigation of functional lateralization. We hypothesized that the inconsistent results in dichotic listening are due to an interaction in processing a mixture of acoustic and linguistic attributes that are differentially processed over the two hemispheres. In this study, a within-subject dichotic listening paradigm was designed, in which different levels of speech and linguistic information was incrementally included in different conditions that required the same tone identification task. A left ear advantage (LEA), in contrast with the commonly found right ear advantage (REA) in dichotic listening, was observed in the hummed tones condition, where only the slow frequency modulation of tones was included. However, when phonemic and lexical information was added in simple vowel tone conditions, the LEA became unstable. Furthermore, ear preference became balanced when phonological and lexical-semantic attributes were included in the consonant-vowel (CV), pseudo-word, and word conditions. Compared with the existing REA results that use complex vowel word tones, a complete pattern emerged gradually shifting from LEA to REA. These results support the hypothesis that an acoustic analysis of suprasegmental information of tones is preferably processed in the right hemisphere, but is influenced by phonological and lexical semantic processes residing in the left hemisphere. The ear preference in dichotic listening depends on the levels of speech and linguistic analysis and preferentially lateralizes across the different hemispheres. That is, the manifestation of functional lateralization depends on the integration of information across the two hemispheres.

1. Introduction

The two hemispheres of the brain are mostly homologous in anatomical structure. But in many cases, high order cognitive function is distributed asymmetrically. One of the most striking examples of this asymmetry was first observed by Paul Broca (1861) and Carl Wernicke (1874) who reported that lesions to the left hemisphere produced deficits in speech and language perception and production. This left hemisphere bias of speech processing was behaviorally observed in neurologically normal people with the seminal dichotic listening studies introduced by Doreen Kimura (1967). This key manipulation in dichotic listening paradigms is to present different auditory information to each ear simultaneously, so that hemispherical lateralization of speech processing can be investigated by assessing the behavioral response bias towards one ear or the other (Hugdahl and Westerhausen, 2016). Classic dichotic listening tasks mostly use English consonant-vowel (CV)

syllables that consist of an initial stop consonant (e.g. /b/, /p/, /g/, /d/, /k/, and /t/) followed by a vowel (e.g. /a/). Participants more frequently report the syllables played to the right ear (e.g. Foundas et al., 2006; Hugdahl et al., 2008). This right ear advantage (REA) is commonly observed during dichotic listening tasks using various English stimuli and provides evidence for a left hemispherical lateralization (Bryden and Murray, 1985; Cutting, 1974; Dwyer et al., 1982; Kimura, 1967; Rimol et al., 2006; Studdert-Kennedy and Shankweiler, 1970).

Different theories have been proposed to account for the REA in dichotic listening. The Structure Model proposed that information from each ear has a stronger representation in the contra-lateral hemisphere due to asymmetry in connecting pathways. Thus, the right ear and its afferent pathway may be more efficient in delivering information to the left hemisphere than the left ear. Together with the left-hemisphere specialization for speech, this leads to a bias in favor of speech stimuli

presented to the right ear (Kimura, 1961a, 1961b). Studdert-Kennedy and Shankweiler. (1970) argued that both hemispheres process auditory parameters of speech signals equally, but the dominant left hemisphere was more efficient in extracting linguistic features from the auditory input. This view was further elaborated by including the influence of top-down processes on extracting auditory information (Kimura, 2011). For example, higher-order cognitive functions (e.g. linguistic and semantic engagement) may modulate lower-order signal perception, which leads to functional asymmetries (Hugdahl and Westerhausen, 2016). Transcranial stimulation studies have provided evidence that the involvement of the left hemisphere leads to the REA bias (D'Anselmo et al., 2015; Prete et al., 2018). Prete and colleagues (2016, 2018, 2019) have provided evidence that in the absence of perceptual stimuli (i.e. auditory imagery imaging a speech), a REA is consistently observed, suggesting that the left hemisphere bias could be accounted for by a left hemisphere specialization in processing speech-related signals.

Interestingly, REA dichotic listening observations are not limited to speech stimuli. For example, the physical features of acoustic stimuli, such as intensity, can modulate ear advantages (Hugdahl et al., 2008). Changes in ear advantages are especially evident when investigating tonal language processing. For example, in Thai, a REA has been observed when listening to consonant-words (i.e. CV syllable words with different consonants but the same tone and vowel) and tone-words (i.e. CV syllable words with different tones but the same consonant and vowel). However, a subgroup of participants showed a left ear advantage (LEA) in a hummed tone condition (i.e. tones with pitch variations when hummed but without any phonological or lexical-semantic information) (Van Lancker and Fromkin, 1973). These mixed dichotic listening results of ear advantage extend to other tonal languages as well, such as Mandarin. Baudoin-Chial, (1986) reported no ear preference whereas Wang and colleagues (Wang et al., 2001) reported a REA for Mandarin tones in native speakers. These mixed results have been observed regardless of whether the tonal speech stimuli were real words (e.g. Ke, 1992; Wang et al., 2004) or pseudo-words (e.g. Shuai and Gong, 2014), and have shown an interaction with intensity between Chinese and English participant groups (Suresh et al., 2017). Dichotic listening results in tonal languages have not show a clear pattern of ear preference in dichotic listening paradigms, in contrast with English studies where consistent REA findings have been reported.

This inconsistency of ear advantage in dichotic listening could be explained by the asymmetric processing of speech across the two hemispheres. The Asymmetric Sampling in Time (AST) theory has proposed that the hemispheric lateralization of speech processing depends on the granularity of analysis (Hickok and Poeppel, 2007; Poeppel, 2003). During the acoustic analysis of temporal and spectral sound features, left and right auditory cortices may have different time constants, where left hemisphere structures process relatively fast frequency information (small temporal constants) and right hemisphere structures bias towards processing low frequency information (large temporal constants) (Poeppel, 2003). These distinct characteristics give rise to the possibility that phonemic and suprasegmental information is preferentially processed over left and right hemispheres respectively. Moreover, lexical access, semantic retrieval, and other higher-level linguistic processes would be processed with a left hemisphere preference (Hickok and Poeppel, 2007). In other words, temporal-spectral features of acoustics signals, as well as the downstream functions in the speech processing hierarchy, drive the differential engagement of the two hemispheres (Belin et al., 2002; Hickok and Poeppel, 2007; Poeppel et al., 2008; Zatorre et al., 2002; Zatorre et al., 1992).

In tonal languages, such as Mandarin, monosyllabic tonal words usually carry at least 3 levels of speech attributes, acoustic (temporal-spectral), phonological and lexical-semantic. For instance, syllables that have the same combination of consonants and vowels can represent different lexical-semantic items if the tones are different – /ma/ with the first tone (flat) means 'mother' and with the third tone (falling-rising) means 'horse'. Under the AST viewpoint, the ear advantage in dichotic

listening should shift as a function of processing attributes of speech stimuli. The commonly found REA in studies using syllables, as well as other types of English stimuli is due to the segmental size of the item, lexical access, and other upstream linguistic processes in the left hemisphere (e.g. Bryden and Murray, 1985; Cutting, 1974; Dwyer et al., 1982). However, when suprasegmental information such as pitch and tones are processed, the right hemisphere would be recruited and dominate performance during dichotic listening, producing a LEA. Indeed, during dichotic listening, processing melodies has shown a LEA whereas processing English syllables has shown a REA (Zatorre, 1979), reflecting a left hemisphere contribution to extracting phonological cues, and a right hemisphere contribution to processing pitch contours (Zatorre et al., 1992).

The current study tests the hypothesis that ear advantages in dichotic listening are due to the functional lateralization along a speech processing hierarchy. More specifically, we investigated whether phonological and lexical-semantic processes can influence tonal processing. In a dichotic listening paradigm with tonal speech sounds that vary in phonological and lexical-semantic levels, a tone identification task was used in all conditions to investigate how the information in different levels of speech processing would influence the ear advantage of tone identification. If the tonal processing in the right hemisphere operates independently from phonological and lexical-semantic processes, the LEA should be evident in all conditions regardless of the level of speech information in the stimuli. Whereas, if the ear advantage is driven by the functional lateralization of speech processing and tonal processing is influenced by phonological and lexical-semantic processes, the results of the tonal identification tasks in dichotic listening should show a LEA for tones that only carries acoustic features and shift to a REA when phonological and lexical-semantic features are added.

2. Methods

2.1. Participants

Twenty-three Mandarin native speakers (7 males; mean age = 20.3, standard deviation = 2.71 years, range 18–27 years) participated in Experiment 1. One participant did not follow the instructions, and one participant did not finish the experiment. Therefore, twenty-one participants were included in the experiment. All participants were right-handed with normal hearing.

Twenty Mandarin native speakers (6 males; mean age = 22.9, standard deviation = 2.32 years, range 19–27 years; right-handed) with normal hearing participated in Experiment 2. All participants were right-handed with normal hearing and did not participate in Experiment 1.

This study was approved by the local IRB committee at New York University Shanghai. Informed written consent was obtained from each participant in both Experiment 1 and 2. Monetary compensation was provided for participation.

2.2. Materials

Four categories of Mandarin tones were used, including hummed tones, simple vowel tones, CV pseudo-word tones, and CV word tones (Fig. 1). Four tones were included in each type of stimulus. Hummed tones were pronounced by recording a native speaker of Mandarin humming the four tones. The vowel /i/ was used in the simple vowel tones. The CV syllable /gi/ was used in pseudo-word tones, and the CV syllable /di/ was used in CV word tones. While tonal features were included across all conditions, each stimulus type varied across phonemic, syllabic and lexical-semantic features (see Table 1). Hummed tones did not include any linguistic features, simple vowel tones included phonemic and lexical-semantic features, CV pseudo-word tones included phonemic and syllabic features, and CV word tones include all features.

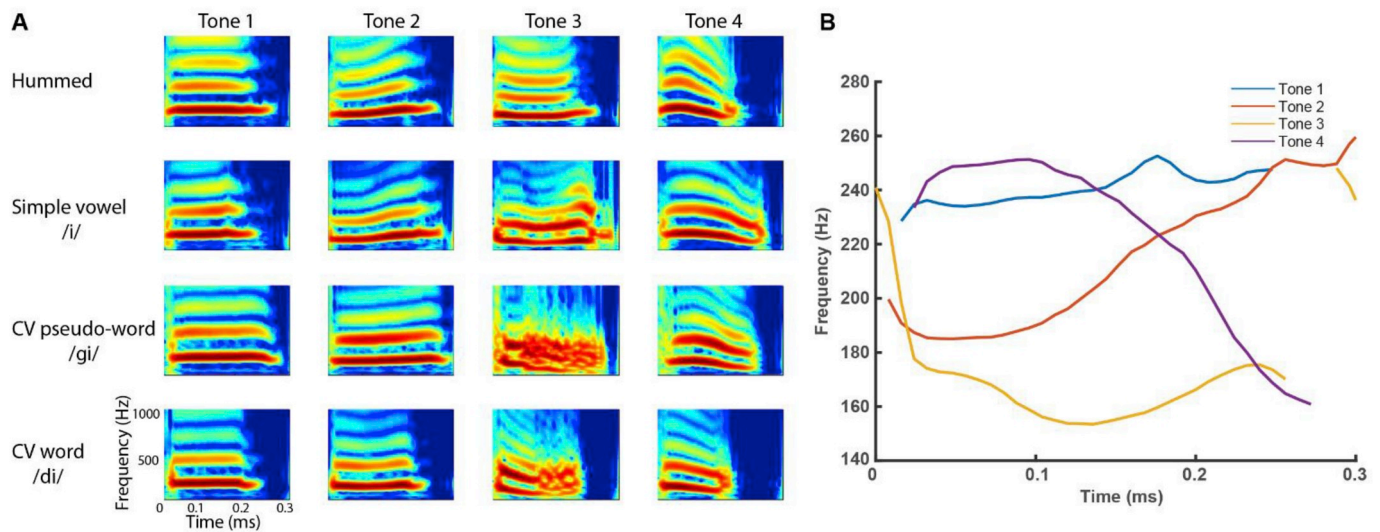Auditory stimuli were recorded in a sound-attenuated room using

Fig. 1. Depiction of stimuli used in the study. A) Spectrograms of all stimuli. Each row includes one type of stimuli and each column includes one of four tones. B) Pitch contour of four tones. Four tones in the simple vowel /i/ are demonstrated as representative samples.

Table 1
Featural components in each type of stimuli. The inclusion of four features components – tone, vowel, consonant, and lexical-semantic—were manipulated in each type of stimuli (hummed tones, simple vowel tones, CV pseudo-word tones, and CV word tones). Tones were included in all types of stimuli. Hummed tones include only tone but no other 3 features. Simple vowel tones include tone as well as vowel /i/ and lexical-semantic features. CV pseudo-word tones include tone feature and vowel /i/ and consonant /g/, but no lexical-semantic feature. CV word tones include all 4 features.

	Tone	Vowel	Consonant	Lexical-semantic
Hummed tones	Yes	No	No	No
Simple vowel (/i/) tones	Yes	Yes	No	Yes
CV pseudo-word (/gi/) tones	Yes	Yes	Yes	No
CV word (/di/) tones	Yes	Yes	Yes	Yes

Shure Beta 58 A microphone. A female speaker pronounced each stimulus 10 times. The continuous auditory signals were recorded (sampled at a rate of 44.1 kHz) and further processed using Praat. One token was selected for each tone of every stimulus type. Hummed tones were 310 ms, simple vowel tones were 322 ms, CV mono-syllabic pseudo-word tones were 342 ms, and CV mono-syllabic word tones were 287 ms. All stimuli were normalized by average intensity (root-mean-square) and delivered through Sennheiser HD280 headphones at about 70 dB SPL. Hummed tones and simple vowel tones were used in Experiment 1, and all four types of stimuli were used in Experiment 2. As shown in Fig. 1, spectrograms were constructed using a log-spaced filter bank approach designed to estimate cochlear critical bands (Flinker et al., 2019) and pitch contours were calculated using a sawtooth waveform inspired pitch estimator (Camacho and Harris, 2008).

2.3. Procedure

Participants first performed an audiometer hearing threshold test. This test quantified whether participants had similar hearing thresholds in both ears. In this test, a single pure tone was pseudo-randomly presented to either left or right ear in each trial. Pure tones in five different frequencies (100 Hz, 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz) were used. The intensity decreased after participants indicated they could hear the pure tone. The relative hearing thresholds between the ears of each participant were calculated by the difference between their left and right ear's thresholds. If the threshold differences between their left and right ears were less than 5 dB at all levels, it indicated participants had

similar hearing thresholds in both ears and could participate in the main experiment.

Next, participants performed the dichotic listening task (Fig. 2). In each trial, a fixation cross was presented at the center of the screen. After 50 ms a pair of different tones was presented to both ears simultaneously. The fixation cross disappeared 50 ms after sound offset and participants were asked which tone they heard first by pressing the corresponding button using keyboard number keys from 1 to 4. Responses were self-paced and participants were encouraged to respond as fast as possible. A number corresponding to the response was presented on the screen for 450 ms after each response. Twelve unique pairs were constructed by pairing two out of four tones of each type of stimuli. Each pair was repeated 15 times, and 180 trials in total were presented in a block. When responding, participants were asked to press 1–4 on the top of the keyboard with their index finger, and the response hand was counterbalanced among participants.

In Experiment 1, hummed tones and simple vowel tones were presented in separate blocks. These two blocks were repeated three times and participants were instructed to perform the dichotic listening tasks in different conditions for each repetition. First, participants were asked to maintain equal attention to both ears during each trial (balanced blocks). Next, two additional conditions (attentional blocks) were run. Participants were instructed to pay attention to either the left or right ear in two blocks (hummed and simple vowel tones) then switch attention to the other ear in the subsequent two blocks. Because strong attentional effects can overwrite ear preference, we used these blocks as a control to verify whether participants were following instructions. The order of blocks in each condition was counterbalanced across participants.

In Experiment 2, all four types of stimuli (hummed tones, simple vowel tones, CV pseudo-word tones, and CV word tones) were used and each type of stimuli was presented in a separate block. A set consisted of a hummed tones block and a simple vowel tones block, and the other set consisted of the other two blocks. The order of blocks in a set and the order of sets were randomized. Participants were asked to maintain equal attention to both ears throughout Experiment 2.

2.4. Data analysis

A *correct* trial is defined as a match of response to the stimulus presented to a given ear. For example, when the stimuli were the first tone to the left ear and the second tone to the right ear, and the participant responded by pressing the key "1", we defined this trial as a *left correct*. If
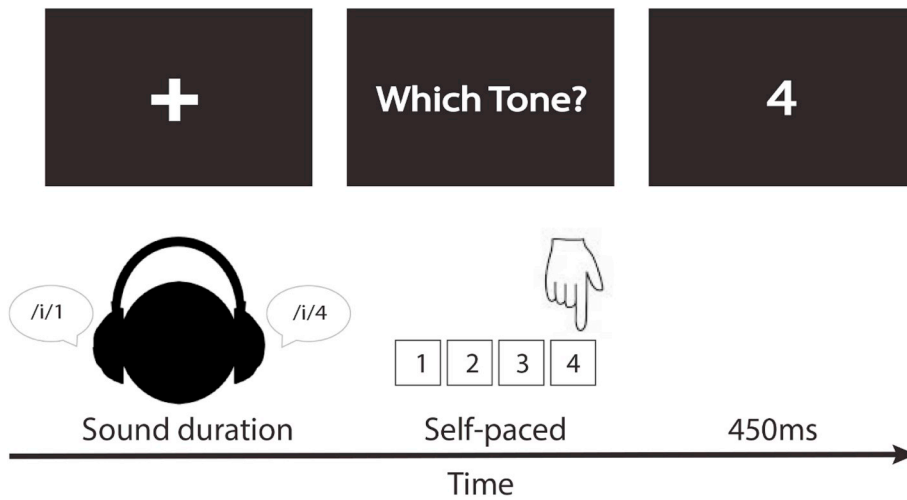
Fig. 2. Depiction of experimental procedure. Participants looked at a fixation in the center of the screen while listening to two different tones presented to their ears. The tones can be in different types of speech stimuli, including hummed tones, simple vowel tones, CV pseudo-word tones, and CV word tones. Here we use simple vowel tones for demonstration. The "/i/1" and "/i/4" represent the first and fourth tones of simple vowel /i/. Participants were asked which tone they heard first, and responded by pressing "1", "2", "3", or "4" key as fast as possible. The response was shown on the screen for 450 ms before the next trial started.

participants responded by pressing the key "2", we defined this trial as a *right correct*. If participants responded "3" or "4", it did not count as a *correct* trial. The more often participants responded to the stimuli presented to the left ear, the higher count of the *left correct*, and vice versa. For our analysis, we first excluded trials that did not match what was played to either the left or the right ear. *Correct Rate* (*CR*) of a given ear (CR_{ear}) is defined as the ratio between the number of *correct* trials for that ear ($correct_{ear}$) and the total number of responses (after exclusion) (Eq. (1)). CR_{left} and CR_{right} were computed separately. The values of CR_{ear} are between 0 and 1, where 1 stands for completely favor to the given ear and 0.5 stands for no preference responses.

$$CR_{ear} = \frac{correct_{ear}}{total\ responses} \tag{1}$$

A *laterality index* (*LI*) is a normalized measure for ear preference and is defined as the difference of CR_{left} minus CR_{right}, divided by the sum of them (Eq. (2)). *LI* ranges from −1 to +1. For example, a positive *LI* indicates a LEA because such *LI* results from a high CR_{left} and a low CR_{right}, suggesting participants responded to the stimuli to their left ear more often than the right ear. We reported both CR_{ear} and *LI* because both measures have been used in the literature (e.g. Tervaniemi and Hugdahl, 2003; Wang et al., 2001), to facilitate comparison with previous studies.

$$LI = \frac{CR_{left} - CR_{right}}{CR_{left} + CR_{right}} \tag{2}$$

For Experiment 1, LEA was expected in the hummed tone condition, and the CR_{left} was expected to be higher than the CR_{right}. Whereas in the simple vowel condition, no LEA was expected so that the CR_{left} was expected to be the same or lower than the CR_{right}. Similar predictions were expected in experiment 2.

The same procedures were applied to prepare for the data analyses for *reaction time* (*RT*) as the dependent variable.

Finally, we assessed the dichotic listening response patterns across all 4 types of stimuli and compared them with the results of complex word (/fan/) tones from the Wang et al. (2001) study. A linear trend analysis was applied to *LI* values of 4 types of stimuli (hummed tones, simple vowel tones, CV pseudo-word tones, and CV word tones). We applied a cross-validation procedure to estimate the linear trend of the relation between the conditions (independent variable) and *LI* (dependent variable). For each cross-validation iteration, we selected one of the 20 subjects and removed this subject's data and used the rest of the 19 subjects' data to fit a linear regression to predict *LI* as a function of condition. This iterative procedure was repeated until all the subjects were removed from the fitting dataset once, thus, we had 20 linear regression functions. We then compare the regression coefficients of the linear regression functions against zero by a one-sample *t*-test.

3. Results

First, we addressed the laterality of processing tones by testing hummed tones. We applied a two-way mixed-effect ANOVA on the *CR* with the factor between-subject factor of *experiment* and within-subject factor of *ear side*. The main effect of *ear side* was significant [$F(1,41) = 11.129$, $p = 0.0018$, $\eta^2 = 0.213$]. However, neither the main effect of *experiment* [$F(1,41) = 0$, $p = 1$, $\eta^2 = 0$] nor the interaction [$F(1,41) = 0.219$, $p = 0.6419$, $\eta^2 = 0.005$] was significant. In Experiment 1 (Fig. 3, left), the post-hoc comparison revealed a significant LEA, where CR_{left} was significantly different from CR_{right} [$t(22) = 3.143$, $p = 0.003$, BH-FDR corrected]. This observation of a LEA was replicated in Experiment 2 (Fig. 3, right) [$t(19) = 3.532$, $p = 0.0011$, BH-FDR corrected]. These consistent results from both experiments suggested a left ear advantage for hummed tone stimuli. Moreover, these LEA results in the balanced blocks contrasted with the results in the attentional blocks. When attention was focused on a given ear, the CR was higher than 0.8,
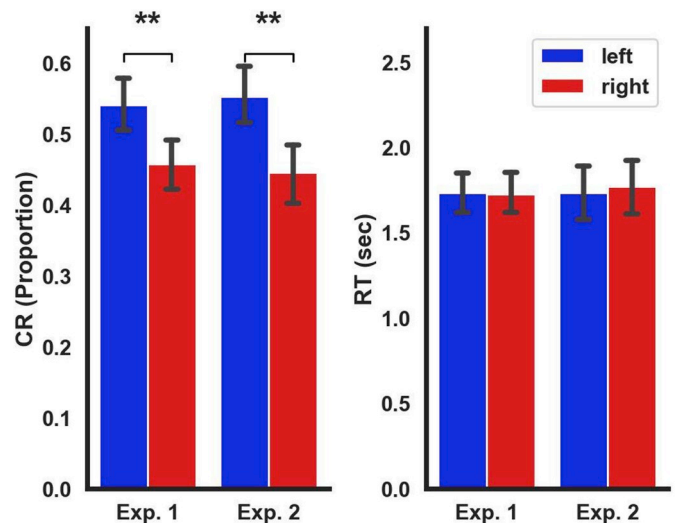

Fig. 3. Results of dichotic listening to the hummed tones. On the left, the *correct rate* (*CR*) is plotted for each ear (left in blue and right in red) in Experiment 1 (left plot) and Experiment 2 (right plot). A robust left ear advantage (LEA) was observed in both experiments. On the right, the reaction time (RT) is plotted for each ear (left in blue and right in red) in Experiment 1 (left plot) and Experiment 2 (right plot). No RT effect was observed. *: $p < 0.05$, **: $p < 0.01$. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

compared with an ear advantage effect of around 0.55 when attention was not focused (equal attention block). Such a large bias demonstrated that participants followed the instructions in all blocks and strong attentional effects could override ear preference. The same statistical analyses were applied on *RT*, but no significant effects were observed [main effect of ear side, $F(1,41) = 0.382$, $p = 0.54$, $\eta^2 = 0.009$, main effect of experiment, $F(1,41) = 0.06$, $p = 0.8077$, $\eta^2 = 0.001$, nor interaction, $F(1,41) = 1.517$, $p = 0.2251$, $\eta^2 = 0.036$].

We then investigated responses along an increasing speech hierarchy where participants also listened to lexical tones of simple vowels, CV pseudo-words, and CV words. Thus, we applied a repeated measure two-way ANOVA on *CR* in Experiment 2 with factors of *condition* and *ear side*. The main effect of side of ears was significant [$F(1,19) = 5.448$, $p = 0.0307$, $\eta^2 = 0.2230$]. However, neither the main effect of condition [$F(3,57) = 0$, $p = 1$, $\eta^2 = 0$] nor the interaction [$F(3,57) = 1.193$, $p = 0.3177$, $\eta^2 = 0.059$] were significant. A post-hoc comparison between the side of ears revealed that CR_{left} was significantly different from CR_{right}, $t(79) = 2.494$, $p = 0.0147$. The LEA v.s REA in the Hummed tone condition was significant, $t(19) = 2.49$, $p = 0.0437$. However, none of the pairs of LEA and REA in other conditions were significant in the post-hoc comparison after BH-FDR correction (Fig. 4). These inconsistent results between experiments suggest that the lateralization for simple vowel tones, that carry semantic information, does not produce a strong LEA. We applied the same analysis on the *RT* but no effects were significant [main effect of condition, $F(3,57) = 1.392$, $p = 0.2604$, $\eta^2 = 0.068$; main effect of side of ear, $F(1,19) = 0.149$, $p = 0.7033$, $\eta^2 = 0.008$; interaction, $F(3,57) = 0.879$, $p = 0.4413$, $\eta^2 = 0.044$].

Finally, we applied a linear trend analysis which revealed a significant decreasing trend of *LI* from hummed to CV word tones ($t(19) = -150.782$, $p < 0.0001$). These results of a continuum of behavioral responses demonstrate a clear shifting pattern of ear advantage in dichotic listening: a LEA while processing hummed tones with no phonological or lexical-semantic information through a REA when both speech and linguistic attributes are added. The comparison of our results with those in Wang et al. (2001) which used CV complex-vowel word tones further confirms the trend of shifting from a LEA to a REA when left hemisphere functions are more strongly engaged. These processes may include a more in-depth phonological analysis for complex vowels, as well as more demanding lexical-semantic processes. Word frequency of the complex vowel (/fan/) tones (M = 0.053, SD = 0.102) is higher than both the simple vowel tones (M = 0.029, SD = 0.057) and CV mono-syllabic word tones (M = 0.036, SD = 0.091) [$\chi^2 (2110) = 6.71$, $p < 0.05$], suggesting a more heavy engagement of the left hemisphere for phonological and lexical-semantic functions that may drive the observation of a REA in dichotic listening in tonal languages (Baudoin-Chial, 1986; Ke, 1992; Shuai and Gong, 2014; Van Lancker and Fromkin, 1973, 1978; Wang et al., 2001, 2004).

4. Discussion

The current study applied a dichotic listening paradigm to investigate the hemispheric asymmetry in processing Mandarin tones. We found that the lateralization of tone processing was influenced by the content of speech. A LEA was consistently observed in two experiments for processing hummed tones that do not have a phonological or lexical-semantic representation, suggesting the right hemisphere was more sensitive to processing signals with only acoustic features in tonal languages. Moreover, with incremental inclusion of phonological and lexical-semantic attributes, the ear preference shifted to a balance for simple vowel tones, CV pseudo-word, and CV word tones, and to a REA for complex word tones. These results support our hypothesis that the ear preference in dichotic listening is driven by hemispheric lateralization of speech functions along the processing hierarchy.

When the slow frequency modulation was the dominant factor in the acoustic signal of the hummed tones, a LEA was consistently observed in both experiments, clearly contrasting with the commonly observed REA
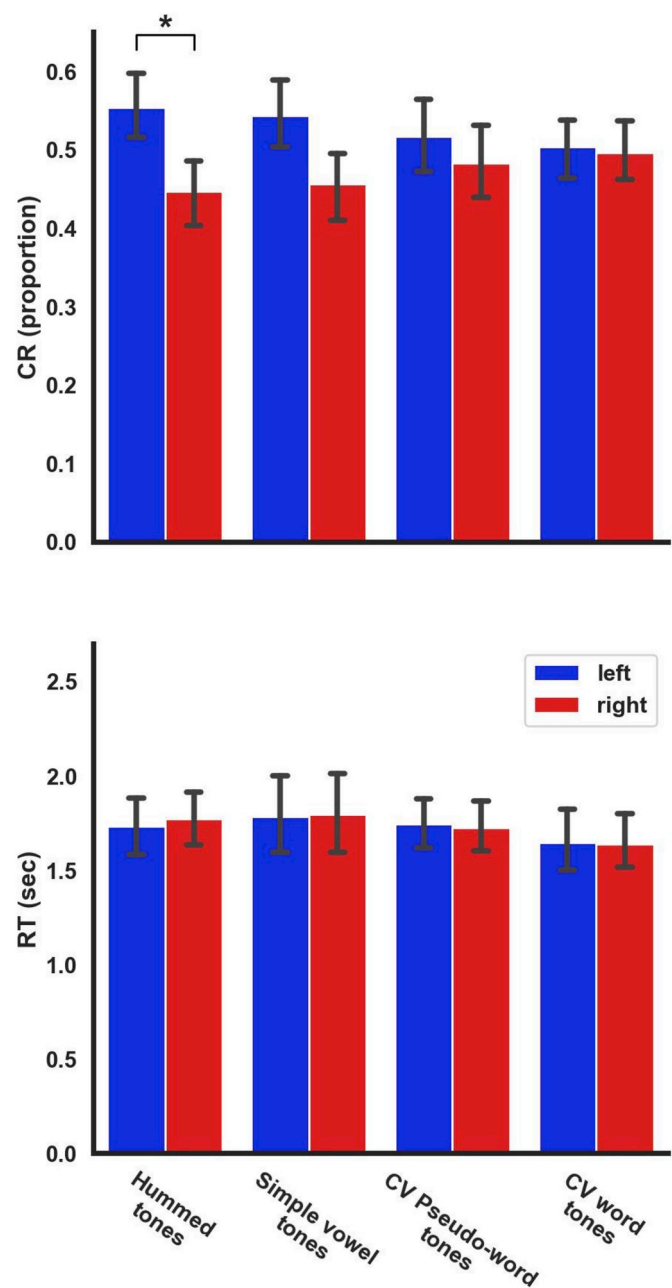

Fig. 4. Results of dichotic listening to four types of tone stimuli in Experiment 2. On the top panel, the *correct rate* (*CR*) was plotted for each ear (left in blue and right in red) for each type of stimuli (hummed tones, simple vowel tones, CV pseudo-word tones, and CV word tones). There was a significant difference between the left CR and right CR. On the bottom panel, the *reaction time* (RT) was plotted for each ear (left in blue and right in red) for each type of stimuli, but no effect was observed. *: $p < 0.05$. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

in dichotic listening of speech (e.g. Bryden and Murray, 1985; Cutting, 1974; Dwyer et al., 1982; Foundas et al., 2006; Kimura, 1967; Rimol et al., 2006; Studdert-Kennedy and Shankweiler, 1970; Tallus et al., 2007). The LEA in dichotic listening of hummed tones demonstrates a right hemisphere lateralization for processing suprasegmental information (with slow temporal characteristics) which is consistent with the model of right lateralization for processing slow auditory modulations (Hickok and Poeppel, 2007; Poeppel, 2003; Zatorre et al., 1992, 2002), low-level acoustic features as the attribute of timbre (e.g. Brancucci and

San Martini, 2003), and loudness discrimination that is mediated by a relatively slow integration (e.g. Brancucci et al., 2005). This robust effect was in line with recent findings in the music domain (e.g. Prete et al., 2015, 2019). A LEA was often observed in pitch perception (Wioland eta l., 1999) and chord recognition (Boucher and Bryden, 1997; Morais et al., 1982), but also see an exception when the task was duration discrimination (Brancucci et al., 2008).

When phonological and lexical-semantic attributes were added to auditory stimuli in the simple vowel tones, CV pseudo-word, and CV word tones conditions, the ear preference in the dichotic listening of tones became more balanced. These results agree with the mixed results in the dichotic listening studies using tonal languages as stimuli (e.g. Gandour et al., 1996; Hugdahl et al., 2008; Jia et al., 2013). That is, phonological and lexical-semantic processes that presumably are dominant over the left hemisphere are active in addition to the process of suprasegmental information that is analyzed preferentially over the right hemisphere. The left hemisphere processing for phonological and lexical-semantic information balance out the preferential right hemisphere processing for suprasegmental information, and hence cause a shift from a clear LEA in the hummed tones conditions to a balanced ear preference in the simple vowel tones, CV pseudo-word, and CV word tones conditions. These ear advantage switching results are also consistent with previous observations of an increased REA advantage in speakers of a second language that was in the same root of their native language (D'Anselmo et al., 2013).

Our results of LI in Experiment 2 further confirmed the trend of shifting from a LEA to a REA when left hemisphere functions are more strongly engaged. Additionally, Wang et al. (2001) used complex CV word tones, and it was predicted to show a stronger shift. Wang et al. (2001) results were in line with our prediction, which was observed in Fig. 5. The Asymmetric Sampling in Time (AST) theory proposed that the hemispheric lateralization of speech processing (Hickok and Poeppel, 2007; Poeppel, 2003). The theory predicts that as the increase of phonological information, the left hemisphere dominance becomes stronger. We observed that a right hemisphere dominance when the acoustic signals contained mainly suprasegmental information (i.e. hummed tones) was influenced by the increasing left hemisphere dominance in processing speech-related features (Poeppel et al., 2008). These results were in line with Hoch and Tillmann (2010) showing that the engagement of both hemispheres in processing phonological and semantic structures in the signal might lead to difficulty in observing a LEA. The gradual decrease of LEA (Fig. 5) suggested that there was a decrease of right hemisphere dominance, including influences from the left hemisphere as opposed to the right hemisphere, presumably a more in-depth phonological analysis for complex vowels, as well as more demanding lexical-semantic processes.

In summary, using a within-subject design dichotic listening paradigm in which we parametrically manipulated the levels of speech and linguistic information, we provide coherent evidence suggesting that the acoustic analysis of suprasegmental information of tones is preferentially processed in the right hemisphere, whereas phonological and lexical-semantic processes weight more towards the left hemisphere. The results of a shift from LEA to REA in dichotic listening as the stimulus attributes incrementally include more suprasegmental, phonological and lexical-semantic information demonstrate that the ear preference in dichotic listening depends on the hemispheric lateralization of function along the speech hierarchy.

CRediT authorship contribution statement

Ning Mei: Conceptualization, Data curation, Formal analysis, Methodology, Writing - original draft, Writing - review & editing. **Adeen Flinker:** Conceptualization, Formal analysis, Methodology, Writing - original draft, Writing - review & editing. **Miaomiao Zhu:** Data curation, Writing - review & editing. **Qing Cai:** Writing - review & editing. **Xing Tian:** Conceptualization, Formal analysis, Methodology, Writing -
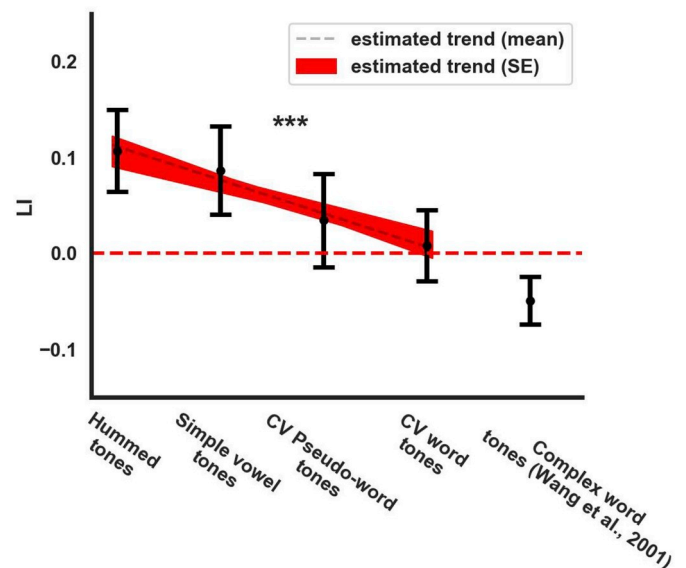

Fig. 5. Linear trend analysis of lateralization index across 4 conditions in experiment 2. Lateralization index (*LI*) was plotted for each of the four types of stimuli (hummed tones, simple vowel tones, CV pseudo-word tones, and CV word tones) in our Experiment 2, as well as the complex word tones in Wang et al. (2001). The black dashed line indicates the average estimate of the linear regression, and the red shaded area indicates the standard error of the estimate of the linear regression. The regressions were fit using only the four types of stimuli in our experiment 2. *** indicates the estimates of the linear regression were significant at the 0.001 level. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

original draft, Writing - review & editing, Funding acquisition, Supervision.

References

Studdert-Kennedy, M., Shankweiler, D., 1970. Hemispheric specialization for speech perception. J. Acoust. Soc. Am. 48 (2B), 579–594. https://doi.org/10.1121/1.1912174.

Baudoin-Chial, S., 1986. Hemispheric lateralization of modern standard Chinese tone processing. J. Neurolinguistics 2 (1–2), 189–199.

Belin, P., Zatorre, R.J., Ahad, P., 2002. Human temporal-lobe response to vocal sounds. Cognit. Brain Res. 13 (1), 17–26. https://doi.org/10.1016/S0926-6410(01)00084-2.

Boucher, R., Bryden, M.P., 1997. Laterality effects in the processing of melody and timbre. Neuropsychologia 35 (11), 1467–1473. https://doi.org/10.1016/s0028-3932(97)00066-3.

Brancucci, A., San Martini, P., 2003. Hemispheric asymmetries in the perception of rapid (timbral) and slow (nontimbral) amplitude fluctuations of complex tones. Neuropsychology 17 (3), 451–457. https://doi.org/10.1037/0894-4105.17.3.451.

Brancucci, A., Babiloni, C., Rossini, P.M., Romani, G.L., 2005. Right hemisphere specialization for intensity discrimination of musical and speech sounds. Neuropsychologia 43 (13), 1916–1923. https://doi.org/10.1016/j.neuropsychologia.2005.03.005.

Brancucci, A., D'Anselmo, A., Martello, F., Tommasi, L., 2008. Left hemisphere specialization for duration discrimination of musical and speech sounds. Neuropsychologia 46 (7), 2013–2019. https://doi.org/10.1016/j.neuropsychologia.2008.01.019.

Bryden, M., Murray, J.E., 1985. Toward a model of dichotic listening performance. Brain Cognit. 4 (3), 241–257. https://doi.org/10.1016/0278-2626(85)90019-3.

Camacho, A., Harris, J.G., 2008. A sawtooth waveform inspired pitch estimator for speech and music. J. Acoust. Soc. Am. 124 (3), 1638–1652. https://doi.org/10.1121/1.2951592.

Cutting, J.E., 1974. Two left-hemisphere mechanisms in speech perception. Percept. Psychophys. 16 (3), 601–612. https://doi.org/10.3758/BF03198592.

D'Anselmo, A., Prete, G., Tommasi, L., Brancucci, A., 2015. The dichotic right ear advantage does not change with transcranial direct current stimulation (tDCS). Brain Stimul. 8 (6), 1238–1240. https://doi.org/10.1016/j.brs.2015.09.007.

D'Anselmo, A., Reiterer, S., Zuccarini, F., Tommasi, L., Brancucci, A., 2013. Hemispheric asymmetries in bilinguals: tongue similarity affects lateralization of second language. Neuropsychologia 51 (7), 1187–1194. https://doi.org/10.1016/j.neuropsychologia.2013.03.016.

Dwyer, J., Blumstein, S.E., Ryalls, J., 1982. The role of duration and rapid temporal processing on the lateral perception of consonants and vowels. Brain Lang. 17 (2), 272–286. https://doi.org/10.1016/0093-934X(82)90021-9.

Flinker, A., Doyle, W.K., Mehta, A.D., Devinsky, O., Poeppel, D., 2019. Spectrotemporal modulation provides a unifying framework for auditory cortical asymmetries. Nat. Human Behav. 3 (4), 393–405. https://doi.org/10.1038/s41562-019-0548-z.

Foundas, A.L., Corey, D.M., Hurley, M.M., Heilman, K.M., 2006. Verbal dichotic listening in right and left-handed adults: laterality effects of directed attention. Cortex 42 (1), 79–86. https://doi.org/10.1016/S0010-9452(08)70324-1.

Gandour, J., Potisuk, S., Ponglorpisit, S., Dechongkit, S., Khunadorn, F., Boongird, P., 1996. Tonal coarticulation in Thai after unilateral brain damage. Brain Lang. 52 (3), 505–535. https://doi.org/10.1006/brln.1996.0027.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8 (5), 393–402. https://doi.org/10.1038/nrn2113.

Hoch, L., Tillmann, B., 2010. Laterality effects for musical structure processing: a dichotic listening study. Neuropsychology 24 (5), 661–666. https://doi.org/10.1037/a0019653.

Hugdahl, K., Westerhausen, R., 2016. Speech processing asymmetry revealed by dichotic listening and functional brain imaging. Neuropsychologia 93, 466–481. https://doi.org/10.1016/j.neuropsychologia.2015.12.011.

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., Hämäläinen, H., 2008. The effect of stimulus intensity on the right ear advantage in dichotic listening. Neurosci. Lett. 431 (1), 90–94. https://doi.org/10.1016/j.neulet.2007.11.046.

Jia, S., Tsang, Y.-K., Huang, J., Chen, H.-C., 2013. Right hemisphere advantage in processing Cantonese level and contour tones: evidence from dichotic listening. Neurosci. Lett. 556, 135–139. https://doi.org/10.1016/j.neulet.2013.10.014.

Ke, C., 1992. Dichotic listening with Chinese and English tasks. J. Psycholinguist. Res. 21 (6), 463–471. https://doi.org/10.1007/BF01067524.

Kimura, D., 1961. Some effects of temporal-lobe damage on auditory perception. Can. J. Psychol. Revue Can. de Psychol. 15 (3), 156–165. https://doi.org/10.1037/h0083218.

Kimura, D., 1961. Cerebral dominance and the perception of verbal stimuli. Can. J. Psychol. Revue Can. de Psychol. 15 (3), 166–171. https://doi.org/10.1037/h0083219.

Kimura, D., 1967. Functional asymmetry of the brain in dichotic listening. Cortex 3 (2), 163–178. https://doi.org/10.1016/S0010-9452(67)80010-8.

Kimura, D., 2011. From ear to brain. Brain Cognit. 76 (2), 214–217. https://doi.org/10.1016/j.bandc.2010.11.009.

Morais, J., Peretz, I., Gudanski, M., 1982. Ear asymmetry for chord recognition in musicians and nonmusicians. Neuropsychologia 20, 351–354. https://doi.org/10.1016/0028-3932(82)90109-9.

Poeppel, D., 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. Speech Commun. 41 (1), 245–255. https://doi.org/10.1016/S0167-6393(02)00107-3.

Poeppel, D., Idsardi, W.J., Van Wassenhove, V., 2008. Speech perception at the interface of neurobiology and linguistics. Phil. Trans. Biol. Sci. 363 (1493), 1071–1086. https://doi.org/10.1098/rstb.2007.2160.

Prete, G., Fabri, M., Foschi, N., Brancucci, A., Tommasi, L., 2015. The "consonance effect" and the hemispheres: a study on a split-brain patient. Laterality 20 (3), 257–269. https://doi.org/10.1080/1357650X.2014.959525.

Prete, G., D'Anselmo, A., Tommasi, L., Brancucci, A., 2018. Modulation of the dichotic right ear advantage during bilateral but not unilateral transcranial random noise stimulation. Brain Cognit. 123, 81–88. https://doi.org/10.1016/j.bandc.2018.03.003.

Prete, G., D'Anselmo, A., Brancucci, A., Tommasi, L., 2019. Unilateral hf-tRNS over the temporal cortex does not influence pleasantness of musical chords. Appl. Acoust. 154, 135–137. https://doi.org/10.1016/j.apacoust.2019.04.035.

Rimol, L.M., Eichele, T., Hugdahl, K., 2006. The effect of voice-onset-time on dichotic listening with consonant–vowel syllables. Neuropsychologia 44 (2), 191–196. https://doi.org/10.1016/j.neuropsychologia.2005.05.006.

Shuai, L., Gong, T., 2014. Temporal relation between top-down and bottom-up processing in lexical tone perception. Front. Behav. Neurosci. 8, 97. https://doi.org/10.3389/fnbeh.2014.00097.

Suresh, C.H., Krishnan, A., Gandour, J.T., 2017. Language experience-dependent advantage in pitch representation in the auditory cortex is limited to favorable signal-to-noise ratios. Hear. Res. 355, 42–53. https://doi.org/10.1016/j.heares.2017.09.006.

Tallus, J., Hugdahl, K., Alho, K., Medvedev, S., Hämäläinen, H., 2007. Interaural intensity difference and ear advantage in listening to dichotic consonant–vowel syllable pairs. Brain Res. 1185, 195–200. https://doi.org/10.1016/j.brainres.2007.09.012.

Tervaniemi, M., Hugdahl, K., 2003. Lateralization of auditory-cortex functions. Brain Res. Rev. 43 (3), 231–246.

Van Lancker, D., Fromkin, V.A., 1973. Hemispheric specialization for pitch and "tone": evidence from Thai. J. Phonetics 1 (2), 101–109. https://doi.org/10.1016/S0095-4470(19)31414-7.

Van Lancker, D., Fromkin, V.A., 1978. Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers. J. Phonetics 6 (1), 19–23. https://doi.org/10.1016/S0095-4470(19)31082-4.

Wang, Y., Jongman, A., Sereno, J.A., 2001. Dichotic perception of Mandarin tones by Chinese and American listeners. Brain Lang. 78 (3), 332–348. https://doi.org/10.1006/brln.2001.2474.

Wang, Y., Behne, D.M., Jongman, A., Sereno, J.A., 2004. The role of linguistic experience in the hemispheric processing of lexical tone. Appl. Psycholinguist. 25 (3), 449–466. https://doi.org/10.1017/S0142716404001213.

Zatorre, R.J., 1979. Recognition of dichotic melodies by musicians and nonmusicians. Neuropsychologia 17 (6), 607–617. https://doi.org/10.1016/0028-3932(79)90035-6.

Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science 256 (5058), 846–849. https://doi.org/10.1126/science.256.5058.846.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cognit. Sci. 6 (1), 37–46. https://doi.org/10.1016/S1364-6613(00)01816-7.