Aging Effects on Mandarin Tone Perception: Evidence from ERP

Authors: Xiao Rong, Liang Dandan, Shanpeng Li, Liang Dandan

Date: 2019-10-10T00:00:00+00:00

Abstract

This study examined the status of pre-attentive tone perception in Mandarin-speaking older adults, investigating the existence of domain-specific aging. Employing event-related potential (ERP) technology and a passive oddball paradigm to elicit MMN responses, the influence of domain-general factors was controlled. The results demonstrated that both tones involving categorical changes and non-speech tones induced attenuated MMN amplitudes, whereas tones not involving categorical changes showed no MMN amplitude decline. The findings indicate that during the pre-attentive stage, domain-specific decline exists in the processing capacity for specific Mandarin tone category knowledge, while perception of tones not involving native phonological knowledge exhibits domain-specific preservation to a certain degree. This preservation is associated with the recruitment of compensatory mechanisms along the temporal dimension. Modulated by compensatory mechanisms, language processing manifests distinct aging trajectories, including decline or preservation.

Full Text

Preamble

Aging Effects on Mandarin Lexical Tone Perception: Evidence from ERPs

XIAO Rong, LIANG Dandan, LI Shanpeng

School of Chinese Language and Culture, Nanjing Normal University, Nanjing 210097, China

Abstract

This study investigated pre-attentive stage lexical tone perception in Mandarin-speaking older adults to examine whether domain-specific aging effects exist. Using event-related potential (ERP) technology and a passive oddball paradigm to elicit mismatch negativity (MMN) responses, we controlled for domain-general factors. The results showed that both lexical tones involving categorical changes and non-speech tones evoked attenuated MMN amplitudes in older adults, whereas tones not involving categorical changes did not show MMN decline. These findings indicate that during the pre-attentive stage, domain-specific decline exists in the processing of specific Mandarin lexical tone category knowledge, while perception of tones not involving native phonemic knowledge shows domain-specific preservation to some extent. This preservation is related to the recruitment of compensatory mechanisms in the temporal dimension. Language processing thus exhibits different aging trajectories of decline or preservation, modulated by compensatory mechanisms.

Keywords: Mandarin Chinese; aging; lexical tone; pre-attention; domain-specific

Research on language processing and aging has focused on the debate between domain-general and domain-specific accounts. The domain-general hypothesis posits that age-related changes in language processing stem from declines in general sensory perception (e.g., auditory function) and general cognitive functions (e.g., working memory, executive function, attention, inhibitory control). In contrast, the domain-specific hypothesis argues that aging effects emerge in specific language cognitive processing mechanisms, with language functions showing particular patterns of decline or preservation (Taylor & Burke, 2002). Specifically, some age-related changes in language-specific functions lead to deteriorated language processing abilities—for instance, weakened connections between phonological and semantic representations result in slower naming speeds (MacKay & Burke, 1990). Conversely, preservation of certain language functions also serves as evidence for domain-specific aging, such as the relative preservation of semantic knowledge compared to significant declines in sensory perception and cognitive function (Persson et al., 2004).

Speech, as the form of language, involves both general sensory-perceptual and cognitive abilities, as well as language-specific processing capacities. Given the complexity of speech, most aging research has focused on minimal segmental phonemes—that is, the perception of individual vowel or consonant phonemes. These studies have extensively debated domain-general versus domain-specific issues (Bidelman, Villafuerte, Moreno, & Alain, 2014). In contrast, suprasegmental phonemes such as lexical tones have received far less attention. Unlike segmental phonemes, which distinguish meaning through timbre, suprasegmental lexical tones systematically manipulate pitch to differentiate meaning (Yip, 2002)—a crucial distinction between tone languages and Germanic languages like English. Mandarin Chinese is a typical tone language, and this study focuses on the aging of lexical tone perception.

1.1 Aging Effects on Lexical Tone Perception: Evidence Supporting Domain-General Accounts

Lexical tone perception involves two levels of information processing: encoding and analysis of acoustic features (primarily frequency) by the peripheral-to-central auditory system, and recruitment of language-related knowledge by the central nervous system. With increasing age, both peripheral auditory sensitivity (Schneider & Pichora-Fuller, 2000) and central auditory function (Lister, Maxfield, Pitt, & Gonzalez, 2011) decline, weakening frequency encoding capabilities. Only three studies have directly examined lexical tone perception in Mandarin-speaking older adults (Wang, Yang, & Liu, 2017; Wang, Yang, & H. Zhang et al., 2017; Yang et al., 2015). These studies all used explicit behavioral identification or discrimination paradigms and provided evidence for domain-general accounts. Wang, Yang, and H. Zhang et al. (2017) found that elevated pitch discrimination thresholds were associated with declines in tone perception ability, suggesting that auditory function deterioration affects tone perception. Kennedy and Raz (2009) found that structural atrophy in the central nervous system impairs general high-level cognitive functions, and such general decline may weaken tone perception, particularly under attentional states that require greater cognitive engagement. Yang et al. (2015) also noted that declines in high-level cognitive functions such as attentional control contribute to reduced tone perception ability.

1.2 Research Methods for Domain-Specific Aging in Lexical Tone Perception: Insights from Segmental Phoneme Studies

No study has directly addressed domain-specific aging in lexical tone perception, but research on segmental phoneme perception and aging provides methodological guidance regarding stimulus comparisons and attentional states.

1.2.1 Two-Level Stimulus Comparisons Examining domain-specific issues in segmental phoneme perception typically involves two levels of comparison. First, speech versus non-speech stimulus processing. EEG studies indicate that speech and non-speech sounds elicit different brain responses in terms of spatial lateralization or amplitude (Shtyrov, Kujala, Palva, Ilmoniemi, & Näätänen, 2000; Sorokin, Alku, & Kujala, 2010). Numerous neuroimaging studies also show that brain activation patterns for speech and non-speech processing are not identical, with language-specific processing regions such as the left planum temporale, bilateral superior temporal sulcus, left inferior frontal gyrus pars opercularis, and ventral premotor cortex. Activation in these regions is not solely related to accessing native phonemic category knowledge—even when passively listening to unfamiliar speech sounds without long-term memory representations, these areas (e.g., left ventral premotor cortex, superior temporal sulcus) are still recruited (Wilson & Iacoboni, 2006), reflecting a broad form of specificity not involving native phonological knowledge processing.

Aging studies have also revealed distinct aging patterns for speech perception compared to non-speech sounds (Du, Buchsbaum, Grady, & Alain, 2016; Geal-Dor, Goldstein, Kamenir, & Babkoff, 2006). Geal-Dor et al. (2006) found in an ERP study that pure tones elicited stronger brain responses than vowels in older adults. Du et al. (2016) reported in an fMRI study that older adults showed greater activation in frontal speech motor regions compared to younger adults only during consonant perception, with no age-group differences in activation levels during non-speech sound perception.

Second, within-phonological-system comparisons of within-category versus between-category changes. Listeners exhibit categorical perception of native phonemes, a phenomenon first discovered by Liberman, Harris, Hoffman, and Griffith (1957). Using a synthesizer to generate a 14-step continuum of C-V syllables varying continuously in formant frequency between /ba/ and /da/, listeners showed higher identification accuracy for changes across category boundaries (i.e., between-category changes) than for equivalent-step changes within the same category (i.e., within-category changes)—a hallmark of categorical perception. This phenomenon originates from native language experience (Durlach & Braida, 1969). Tone language listeners also demonstrate categorical perception when processing native tones, comparing heard tones with stored long-term memory representations to categorize continuously varying sounds into specific tone categories (Qi & Liu, 2015).

Segmental phoneme aging studies provide evidence for age-related changes in categorical processing. Behavioral research shows that aging reduces the degree of categorical perception for consonant phonemes (Gordon-Salant & Fitzgibbons, 1993; Harkrider, Plyler, & Hedrick, 2005). Bidelman et al. (2014) also found in an ERP study that vowel category boundaries were less distinct in older adults compared to younger adults.

1.2.2 Selection of Attentional and Pre-attentive States It is important to note that all the aforementioned speech perception aging studies were conducted under attentional conditions. Only three studies have examined age-related changes in pre-attentive segmental phoneme (consonant) perception (Aerts, Mierlo, & Hartsuiker, 2013; Bellis, Nicol, & Kraus, 2000; Cheng, Baillet, Hsiao, & Lin, 2015), and these have discussed domain-specific perspectives in interpreting their results. Aerts et al. (2013) proposed that each phonemic contrast ages in a different, phoneme-specific manner. The decline in MMN (mismatch negativity) elicited by consonant contrasts may result from reduced "hard-wiring" strength in neuronal ensembles representing consonant oppositions.

Unlike attentional states, pre-attentive states minimize recruitment of high-level cognitive functions, thereby reducing interactive effects of their decline and better enabling investigation of whether more implicit linguistic knowledge is affected by aging (Aerts et al., 2013; Cheng et al., 2015).

1.3 Mechanisms of Phoneme Perception Aging: Decline and Preservation

Phoneme perception abilities exhibit different aging trajectories of decline or preservation, manifesting at both comparison levels described in Section 1.2.1. Some studies find that speech perception abilities decline significantly compared to non-speech sounds in older adults (Geal-Dor et al., 2006), while others report a degree of preservation (Du et al., 2016). Between-category change detection abilities show significant decline compared to within-category changes (Bidelman et al., 2014), yet some studies demonstrate preserved between-category change detection (Bellis et al., 2000). Although decline and preservation appear contradictory, aging is accompanied by self-organization processes in the nervous system, with preservation being related to the recruitment of compensatory mechanisms in older adults.

Compensatory mechanisms underlying preserved phoneme perception abilities manifest in two ways. Spatially, fMRI and other neuroimaging studies reveal compensation through activation of broader brain regions (Du et al., 2016; Getzmann & Falkenstein, 2011). Du et al. (2016) found that during consonant perception, older adults showed greater activation in frontal speech motor regions than younger adults, and this activation correlated with improved task performance in older adults. Temporally, previous research suggests that longer processing time also serves as a compensatory mechanism for perceptual decline (Bellis et al., 2000). Studies have found that while the amplitude of MMN responses to consonant changes in older adults did not differ from younger adults, the peak latency was longer, indicating a temporal compensation mechanism (Bellis et al., 2000).

1.4 Research Questions

Current research on speech processing and aging has focused primarily on segmental phonemes, yet lexical tones differ from segmental phonemes in both natural attributes and linguistic functions. Therefore, the cognitive processing of lexical tone perception should differ from that of segmental phonemes, and its aging patterns may also differ. Moreover, existing studies on lexical tone perception and aging have only examined the role of domain-general factors (Wang et al., 2017; Wang, Yang, & H. Zhang et al., 2017; Yang et al., 2015). However, as a linguistic component, the importance and specificity of lexical tones are more evident in their linguistic functions, and their perception involves language-level processing (Shtyrov et al., 2000; Sorokin et al., 2010; Xi, L. Zhang, Shu, Y. Zhang, & Li, 2010; Yu et al., 2017).

Based on the limitations and insights from previous research, the present study aimed to investigate the neural-level performance of Mandarin-speaking older adults in perceiving lexical tones—a suprasegmental phoneme—under pre-attentive conditions, observing decline and preservation in tone perception abilities to examine whether domain-specific aging exists.

Methodologically, we controlled for general factors by employing a passive oddball paradigm to elicit MMN, an effective electrophysiological marker for assessing automatic central auditory difference detection in pre-attentive stages that does not require conscious stimulus judgment (Näätänen, Paavilainen, Rinne, & Alho, 2007). This paradigm minimizes the influence of high-level cognitive functions on perception. Additionally, we adjusted the sensation level of presented stimuli according to individual auditory threshold declines in older adults to control for effects of peripheral auditory function deterioration (following Ross, Fujioka, Tremblay, & Picton, 2007).

Beyond these factors, central auditory function decline also represents a general form of deterioration affecting tone perception. MMN can reflect both central auditory system detection of acoustic differences and cortical detection of linguistic feature differences (Näätänen et al., 2007). Domain-specificity emerges in contrast to domain-generality. To separate the effects of linguistic function decline from general central auditory deterioration, and to examine the two levels of language processing aging, we introduced control conditions including nonspeech tones, between-category tone changes, and within-category tone changes, comparing MMN responses elicited by different stimulus types.

2. Methods

2.1 Participants

We recruited 18 young adults (mean age = 24.2 years, range = 22.7–29.0 years, 9 males and 9 females) and 22 older adults (mean age = 66.2 years, range = 55.6–79.6 years, 11 males and 11 females). All participants met the following criteria: native Mandarin Chinese speakers with standard Mandarin pronunciation; no history of speech or language developmental disorders; right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971); bachelor' s degree or higher education; normal or corrected-to-normal vision; less than 3 years of musical training; no history of neurological or psychiatric disorders; and not taking medications affecting the nervous system during the experiment. Additionally, older adults scored 27 on the Mini-Mental State Examination (M. F. Folstein, S. E. Folstein, & McHugh, 1975), indicating no dementia.

2.2 Hearing Assessment

We measured participants' bilateral air-conduction pure-tone thresholds at seven frequency intervals from 125–8000 Hz using an audiometer to assess peripheral hearing. All participants had pure-tone average thresholds of 512–4000 Hz 45 dB HL and 6000–8000 Hz 70 dB HL, with no interaural asymmetry (threshold differences >15 dB HL at no more than one frequency), indicating no conductive hearing loss.

2.3 Stimuli and Procedure

2.3.1 Stimuli Stimuli in the EEG experiment included two pairs of speech sounds (between-category contrast, within-category contrast) and one pair of non-speech sounds. We first determined between-category and within-category contrast stimuli through a classic lexical tone categorical perception test. We selected Tone 2 and Tone 4 from Mandarin Chinese as test items because they are the most easily discriminable. Since Chinese syllables correspond to morphemes with semantic functions, and semantic changes can also elicit MMN (Shtyrov & Pulvermüller, 2002), we used pseudo-syllables to exclude semantic interference. We listed all possible Tone 2/Tone 4 pseudo-syllable pairs and had 15 Mandarin speakers rate their semantic association strength (allowing associations from dialects and foreign languages), selecting the 15 pairs with lowest average ratings. A native Mandarin speaker recorded these syllable pairs at a 44.1 kHz sampling rate.

For non-speech stimulus generation, we controlled physical characteristics of speech and non-speech stimuli to investigate speech-specific decline. Research shows that pure tones with lower spectral complexity elicit significantly smaller MMN amplitudes than harmonically rich complex sounds (Takegata, Tervaniemi, Alku, Ylinen, & Näätänen, 2008). To exclude spectral complexity effects, we generated non-speech sounds by rotating speech spectra using a MATLAB script (version R2011a, Mathworks) published by Scott, Blank, Rosen, and Wise (2000; available at http://www.phon.ucl.ac.uk/resource/software.php), which transforms high frequencies to low frequencies and vice versa. However, previous research found that fricatives, nasals, and stops remain similar after rotation (Blesser, 1972), still sounding like speech. To ensure rotated sounds were perceived as non-speech, we had 15 Mandarin speakers rate the speech similarity (i.e., degree of similarity to human-produced speech) of rotated speech pairs. The pair with lowest average speech similarity was the rotated version of /we2/ and /we4/, which we selected as experimental stimuli.

Based on the original recordings, we generated a tone continuum (a set of synthesized signals with continuously varying fundamental frequency). Previous research consistently shows that pitch contour plays the primary role in Mandarin tone perception (Lin, Yan, & Sun, 1984). Therefore, we manipulated tone by modifying syllable pitch contours using Praat software. To control for acoustic features other than pitch contour (voice onset time, average intensity, duration, pitch, envelope, etc.), we removed the pitch contour from /we2/ while preserving other spectral components, then copied the pitch contour from /we4/ onto the /we2/ spectrum, creating two stimuli with 250 ms duration (including 10 ms rise/fall time) at 70 dB intensity.

We used these two stimuli as endpoints of the tone continuum and employed PSOLA synthesis to generate 15 stimuli for the categorical perception test. The pitch contours are shown in Figure 1 [Figure 1: see original paper]. Participants sat before a screen, read instructions, and pressed any key to begin. A practice block of 10 trials preceded the formal experiment, requiring participants to identify whether the sound played through computer speakers was Tone 2 or Tone 4 via button press. Each stimulus was presented randomly in isolation 20 times. The test lasted approximately 10 minutes. We calculated the proportion of Tone 4 identifications for each step of the continuum in both young and older groups, generating categorical perception curves shown in Figure 2 [Figure 2: see original paper].

As shown in Figure 2, T5 and T10 represent the inflection points of categorical perception for both groups, with tones at T5 and earlier most likely identified as Tone 2, and tones at T10 and later as Tone 4. In the EEG experiment, the between-category tone contrast block used T10 as the standard stimulus and T5 as the deviant, differing in both pitch contour and tone category. The acoustic difference magnitude between standard and deviant stimuli was identical for between-category and within-category contrasts. The within-category tone contrast block used T10 as standard and T15 as deviant, differing only in pitch contour. Before the EEG experiment, all participants again identified which of the four tones each of the three stimuli represented. Results confirmed that all participants identified T5 and T10 as different tone categories and T10 and T15 as the same category, establishing that T10-T5 represents a between-category change and T10-T15 a within-category change.

We removed pitch contours from speech stimuli using Praat, then applied spectral rotation to generate non-speech stimuli without lexical tones. We then copied pitch contours from T10 and T15 onto the non-speech spectra to create standard and deviant stimuli for the non-speech tone contrast block, ensuring identical pitch contour differences between deviant and standard as in the within-category tone contrast block.

2.3.2 Experimental Procedure All EEG experiments used a classic auditory passive oddball paradigm. Standard stimuli occurred with 90% probability and deviant stimuli with 10% probability. Three blocks with identical procedures were administered: between-category tone contrast, within-category tone contrast, and non-speech tone contrast. Each block contained 1015 trials, with the first 15 stimuli being standards. Deviant stimuli were presented in pseudo-random order with at least two standards between successive deviants. The inter-stimulus interval was 650 ms. The three blocks were presented in random order, each lasting 15 minutes with 5-minute breaks between blocks.

Participants were instructed to watch a self-selected silent video and ignore sounds presented through headphones. The video computer was positioned 70 cm from participants. Stimuli were presented binaurally at 70 dB for all young participants. For older adults, presentation intensity was adjusted based on individual air-conduction hearing thresholds: for every 5 dB that an older adult' s threshold exceeded the young group' s average threshold (X), intensity was increased by 5 dB. This controlled the sensation level (SL) across groups, excluding the possibility that ERP effects simply reflected hearing differences.

2.3.3 EEG Recording

EEG data were recorded using a Brain Products SynAmps2 amplifier with a 32-channel electrode cap, referenced to left and right mastoids. AC sampling was employed at 500 Hz/channel, with scalp impedance maintained below 5 kΩ.

2.3.4 ERP Data Analysis

Offline processing used Brain Vision Analyzer 2.1 software. The analysis epoch spanned from 100 ms pre-stimulus to 500 ms post-stimulus (600 ms total). Blink artifacts were automatically corrected, and EEG events with amplitudes exceeding ±100 V were automatically rejected as artifacts. Offline filtering applied a 1–50 Hz bandpass. EEG signals were then averaged separately for each experimental condition.

Two participants from each group were excluded due to excessive EEG artifacts, leaving data from 20 older and 16 young participants for analysis. Figure 3 [Figure 3: see original paper] shows ERP responses at electrode Fz for both groups across the three tone conditions.

The grand-averaged MMN waveform at electrode Fz peaked at 226 ms post-stimulus. We defined the analysis time window as 196–256 ms (30 ms on either side of the peak). First, we conducted reliability tests for MMN using one-sample t-tests comparing MMN amplitudes at electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4 against zero. Results showed reliable MMN at all electrodes in the young group across all conditions ($ps < 0.05$). In the older group, reliable MMN was elicited at all electrodes except P3, Pz, and P4 in the non-speech tone contrast condition ($ps > 0.05$), and at P4 in the within-category tone contrast condition where MMN amplitude was marginally significant ($p = 0.062$; all other electrodes $ps < 0.05$).

We performed separate 3 (tone type: between-category/within-category/non-speech) × 2 (group: older/young) repeated-measures ANOVAs on MMN peak latency and mean amplitude elicited by deviant stimuli. Results are shown in Table 1 .

Table 1 Results of ANOVA with tone type and group as independent variables and peak latency/mean amplitude as dependent variables

Effect	df	Peak Latency F	Mean Amplitude F
Tone Type	(1,34)	3.238†	–
Group	(1,34)	14.201**	4.973*
Group × Tone Type	(1,34)	3.346*	8.224**

Note: ** $p < 0.01$, * p 0.05, † $0.05 < p < 0.07$, all other ps 0.07.

Simple effects analyses for significant interactions revealed: For latency, older adults showed significantly longer MMN peak latencies than young adults for between-category tones ($p < 0.001$) and within-category tones ($p = 0.027$), but no significant group difference for non-speech tones ($p = 0.780$), as shown in Figure 4 [Figure 4: see original paper].

For amplitude, older adults showed significantly reduced MMN amplitude compared to young adults for between-category tones ($p = 0.02$) and non-speech tones ($p = 0.002$), but no group difference for within-category tones ($p = 0.448$), as shown in Figure 5 [Figure 5: see original paper].

Finally, we conducted a 2 (laterality: left/right) × 3 (location: frontal/central/temporal) × 2 (group: young/older) repeated-measures ANOVA on mean MMN amplitude across the three tone contrast conditions. Electrodes F3, F4, C3, C4, P3, P4 were selected for analysis within the 196–256 ms time window (matching the amplitude-maximal Fz window; reference Ren, Tang, Li, & Sui, 2013). Results are shown in Table 2 .

Table 2 Results of ANOVA with location, laterality, and group as independent variables and mean amplitude as dependent variable

Effect	df	F
Location	(2,68)	8.446**
Group	(1,34)	5.625*
Laterality	(1,34)	9.394**
Laterality × Location	(2,68)	15.866**
Laterality × Group	(1,34)	22.073**
Group × Location	(2,68)	3.864*
Laterality × Location × Group	(2,68)	–

Note: ** $p < 0.01$, * p 0.05, † $0.05 < p < 0.07$, all other ps 0.07.

These results showed significant main effects of group only in the between-category and non-speech contrast conditions, with older adults showing significantly smaller mean MMN amplitude than young adults. Main effects of location were significant across all three conditions. Post-hoc comparisons revealed that frontal and central regions showed significantly greater amplitude than temporal regions (ps < 0.01), with no significant differences between frontal and central regions (ps > 0.05).

In the non-speech condition, the laterality × location interaction was significant. Simple effects analysis showed that in the left hemisphere, frontal and central regions were significantly greater than temporal regions (ps < 0.01), with no frontal-central difference ($p = 0.34$). In the right hemisphere, all three regions differed significantly from each other, with amplitude decreasing from frontal > central > temporal (ps < 0.05).

Given the wide age range in our older group (55.6–79.6 years) and potential heterogeneity, we conducted supplementary analyses on the older group alone, adding age as a covariate. These analyses differed from previous analyses in two ways: (1) age was included as a covariate, and (2) analyses were restricted to the older group. Results showed that in the between-category contrast condition, the age × laterality interaction was significant ($p = 0.013$, $^{2} = 0.295$), with lateralization shifting from left to right with increasing age. No other analyses showed significant main effects of age or age × factor interactions ($ps > 0.05$).

4. Discussion

This study investigated pre-attentive lexical tone perception in Mandarin-speaking older adults, focusing on the role of domain-specific factors in language processing aging. Results showed that in the between-category condition, older adults exhibited consistent and significant decline compared to young adults in both MMN peak latency and mean amplitude (i.e., reduced amplitude, prolonged latency). In the non-speech condition, only mean amplitude showed significant decline. In the within-category condition, only peak latency showed significant prolongation.

MMN amplitude is generally considered more reliable than latency (Zhao, 2010); therefore, our conclusions primarily rely on statistical results using mean amplitude as the dependent variable, supplemented by peak latency results. Our findings demonstrate that even during the low cognitive-demand pre-attentive stage, Mandarin-speaking older adults show age-related decline in perceiving native tones involving categorical changes. In contrast, perception of tones not involving native phonemic knowledge shows relative preservation compared to general central auditory system decline in pitch change detection. This preservation likely originates from recruitment of compensatory mechanisms in the temporal dimension. Our study provides new evidence from Mandarin Chinese for the central debate in language processing aging research and reveals neural-level patterns of suprasegmental phoneme perception aging.

4.1 Domain-Specific Preservation in Speech Versus Non-speech Comparisons

This comparison examined MMN responses to within-category tones and non-speech tones with identical physical changes and no phonological category changes, allowing investigation of domain-specific aging at the broad speech versus non-speech level. Results showed that older adults had significantly smaller MMN mean amplitude for non-speech tone contrasts compared to young adults, while no group difference emerged for within-category tone discrimination. Within the young group, non-speech tone contrasts elicited larger MMN mean amplitude than speech tones sharing the same pitch contour, an enhancement effect absent in the older group.

Geal-Dor et al. (2006) examined ERP responses to non-speech sounds, vowels, and other signals in older adults, concluding that speech perception aging is domain-specific because brain responses to speech differed from non-speech.

However, their non-speech stimuli had much simpler spectral structure than speech stimuli, making differences potentially attributable to declines in primary acoustic signal analysis. Our study controlled spectral complexity by generating non-speech through spectral rotation, making early physical feature processing similar for speech and non-speech (Christmann, Berti, Steinbrink, & Lachmann, 2014). Nevertheless, older adults showed different aging patterns: clear decline in non-speech tone discrimination but domain-specific preservation in speech tone discrimination not involving phonemic category knowledge. Research shows that language functions are relatively preserved compared to general sensory perception and high-level cognition in older adults (Persson et al., 2004), supporting domain-specific accounts of language processing aging.

In summary, Mandarin speakers exhibit domain-specific aging patterns for speech tone perception not involving native phonemic knowledge, contrasting with non-speech tones that reflect general central auditory processing decline. This indicates that while general central auditory processing of pitch shows significant age-related decline, speech tone perception shows domain-specific preservation.

4.2 Domain-Specific Decline in Between-Category Versus Within-Category Comparisons

This comparison examined MMN responses to between-category and within-category tones with identical acoustic change magnitude but only the former involving category changes, allowing investigation of aging effects in category knowledge recruitment. Results showed that older adults had significantly smaller response amplitude for between-category tone contrasts than young adults, with no group difference for within-category contrasts. Within the young group, between-category contrasts elicited larger MMN mean amplitude than within-category contrasts, a category-change enhancement effect absent in older adults.

These findings indicate that Mandarin-speaking older adults cannot process lexical tone category information at the pre-attentive stage. In contrast, young adults automatically extract category information when processing speech signals involving native phonemic contrasts while showing reduced sensitivity to within-category phonetic differences. This demonstrates age-related decline in tone category information recruitment during pre-attentive processing.

The three existing studies examining pre-attentive phoneme perception aging yielded inconsistent conclusions (Aerts et al., 2013; Bellis et al., 2000; Cheng et al., 2015). Early research found no age-group differences in MMN amplitude or latency for consonant contrasts, concluding no pre-attentive decline in phoneme discrimination (Bellis et al., 2000). Unlike our study which used frontal electrode Fz for analysis, Bellis et al. (2000) used temporal electrodes. Previous research shows frontal electrodes elicit more reliable MMN than temporal sites (Xi et al., 2010). Additionally, our older participants averaged 66.2 years, whereas

Bellis et al.' s (2000) participants were around 55 years—an age at which neural representations of phonemic changes may not yet show significant decline. Our results align with Aerts et al. (2013), who found prolonged MMN latency and reduced amplitude for between-category consonant contrasts in older adults at Fz and Cz. Cheng et al. (2015) using MEG also found increased MMNm latency and reduced activation in older adults, both studies concluding age-related decline in pre-attentive phoneme discrimination.

However, these studies only included between-category (phonemic-level) stimuli, meaning deviants differed from standards in both acoustic features and phonemic categories. Consequently, age effects could be attributed to general auditory decline. Our study demonstrates domain-specific decline in categorical processing by comparing MMN differences between within-category and between-category tone contrasts. Research suggests that pre-attentive discrimination of different speech units relates to activation of different long-term speech memory traces (Näätänen et al., 1997). These traces, based on patterns of phonetic contrast identification, mature with language development and are specific to one' s native language (Shtyrov et al., 2000). Long-term memory traces for tone categories may undergo age-related changes, with weakened internal connection strength (Aerts et al., 2013), leading to decline in tone category discrimination ability.

4.3 Domain-Specific Preservation Related to Compensatory Mechanism Recruitment

As noted, our study found domain-specific preservation in older adults' within-category tone discrimination ability. Previous research suggests that increased processing duration reflects compensatory mechanism recruitment in the temporal dimension during aging (Bellis et al., 2000). Our peak latency results across the three conditions were as follows: no group difference in MMN peak latency for non-speech tones; significantly longer MMN peak latency in older adults for within-category tones; and similarly prolonged latency for between-category tones. Combined with amplitude comparisons, we speculate that older adults did not recruit compensatory mechanisms for non-speech tones; compensation occurred only at the speech level but with differential effectiveness—insufficient for between-category conditions but adequate for within-category conditions.

This difference reflects varying aging trajectories within language processing. Initially, compensation may offset decline, resulting in preservation, but for some aspects, compensatory recruitment cannot sufficiently counteract greater age-related deterioration. In our study, older adults showed greater decline in perceiving speech involving category changes compared to broad speech without phonemic category changes. Although peak latency increased, compensation failed to reach adequate levels, resulting in obvious decline.

Spatially, previous research found that neural recruitment specificity decreases with age during language processing, with older adults recruiting broader spatial regions, interpreted as a compensatory mechanism (Dennis & Cabeza, 2008). However, our spatial distribution analyses revealed no significant interactions between group and laterality (left/right), group and location (frontal/central/temporal), or group × laterality × location for mean MMN amplitude across all three tone change conditions. This indicates no age-related changes in MMN spatial distribution. Across all conditions, both groups showed no significant frontal-central differences, with both regions significantly greater than temporal regions. This aligns with previous pre-attentive consonant perception aging research (Aerts et al., 2013). Consistent findings show MMN primarily distributes across frontal and central regions (Näätänen et al., 1997; Xi et al., 2010). Additionally, no hemispheric lateralization effects emerged for MMN amplitude in either group across conditions. Cheng et al. (2015) using higher spatial resolution MEG also found no age-related changes in lateralization patterns for mismatch responses to consonant contrasts. Our and previous results indicate that aging does not universally reduce spatial distribution specificity; MMN responses show stable, age- and stimulus-resistant spatial characteristics of frontal/central and bilateral distribution. Thus, no spatial compensatory mechanism exists for MMN responses.

One limitation is our older group' s wide age range (55.6–79.6 years) and substantial heterogeneity. Supplementary analyses with age as a covariate revealed a significant age × laterality interaction only in the between-category contrast condition, with lateralization shifting from left to right with increasing age. No other analyses showed significant age main effects or age × factor interactions, indicating that age directly affects neural mechanisms of tone perception to some degree when processing tones involving native phonemic category knowledge. Notably, the young group (22.7–29.0 years) showed no lateralization effects (bilateral symmetric activation), suggesting potential changes in lateralization emergence across the lifespan. Future research should subdivide older adults into narrower age bands and include middle-aged groups to further examine trajectories of tone perception aging.

Conclusion

Pre-attentive lexical tone perception in Mandarin-speaking older adults shows language domain-specific decline and preservation: domain-specific decline in processing specific Mandarin tone category knowledge, and domain-specific preservation in broader speech tone processing abilities. Preservation and decline reflect different aging stages modulated by compensatory mechanisms, demonstrating that language processing aging is domain-specific.

References

Aerts, A., van, Mierlo, P., & Hartsuiker, R. J. (2013). Neurophysiological investigation of phonological input: Aging effects and development of normative data. *Brain and Language, 125*(3), 253–263.

Bellis, T. J., Nicol, T., & Kraus, N. (2000). Aging affects hemispheric asymmetry in the neural representation of speech sounds. *Journal of Neuroscience, 20*(2), 791–797.

Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014). Age-related changes in the subcortical–cortical encoding and categorical perception of speech. *Neurobiology of Aging, 35*(11), 2526–2540.

Blesser, B. (1972). Speech perception under conditions of spectral transformation: I. Phonetic characteristics. *Journal of Speech, Language, and Hearing Research, 15*(1), 5–41.

Cheng, C. H., Baillet, S., Hsiao, F. J., & Lin, Y. Y. (2015). Effects of aging on the neuromagnetic mismatch detection to speech sounds. *Biological Psychology, 104*, 48–55.

Christmann, C. A., Berti, S., Steinbrink, C., & Lachmann, T. (2014). Differences in sensory processing of German vowels and physically matched non-speech sounds as revealed by the mismatch negativity (MMN) of the human event-related brain potential (ERP). *Brain and Language, 136*, 8–18.

Dennis, N. A., & Cabeza, R. (2008). Neuroimaging of healthy cognitive aging. *The Handbook of Aging and Cognition, 3*, 1–54.

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature Communications, 7*, 1–12.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception. I. Preliminary theory of intensity resolution. *The Journal of the Acoustical Society of America, 46*(2B), 372–383.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state" : A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198.

Geal-Dor, M., Goldstein, A., Kamenir, Y., & Babkoff, H. (2006). The effect of aging on event-related potentials and behavioral responses: Comparison of tonal, phonologic and semantic targets. *Clinical Neurophysiology, 117*(9), 1974–1989.

Getzmann, S., & Falkenstein, M. (2011). Understanding of spoken language under challenging listening conditions in younger and older listeners: A combined behavioral and electrophysiological study. *Brain Research, 1415*, 8–22.

Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech, Language, and Hearing Research, 36*(6), 1276–1285.

Harkrider, A. W., Plyler, P. N., & Hedrick, M. S. (2005). Effects of age and spectral shaping on perception and neural representation of stop consonant stimuli. *Clinical Neurophysiology, 116*(9), 2153–2164.

Kennedy, K. M., & Raz, N. (2009). Aging white matter and cognition: Differential effects of regional variations in diffusion properties on memory, executive functions, and speed. *Neuropsychologia, 47*(3), 916–927.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*(5), 358.

Lin, M. C., Yan, J. Z., & Sun, G. H. (1984, February). A preliminary experiment on the normal stress of two-character Chinese. *Dialect,* (1)*, 57–73.

Lister, J. J., Maxfield, N. D., Pitt, G. J., & Gonzalez, V. B. (2011). Auditory evoked response to gaps in noise: Older adults. *International Journal of Audiology, 50*, 211–225.

MacKay, D. G., & Burke, D. M. (1990). Chapter five cognition and aging: A theory of new learning and the use of old connections. *In Advances in Psychology* (Vol. 71, pp. 213–263). North-Holland.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432–434.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*, 2544–2590.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Persson, J., Sylvester, C. Y. C., Nelson, J. K., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2004). Selection requirements during verb generation: Differential recruitment in older and younger adults. *Neuroimage, 23*(4), 1382–1390.

Qi, B. E., & Liu, B. (2015). The review of categorization features of tone perception. *Journal of Clinical Otorhinolaryngology Head and Neck Surgery, 29*(15), 1396–1400.

Ren, G. Q., Tang, Y. Y., Li, X. Q., & Sui, X. (2013). Pre-attentive processing of Mandarin tone and intonation: Evidence from event-related potentials. *In Functional brain mapping and the endeavor to understand the working brain*. IntechOpen.

Ross, B., Fujioka, T., Tremblay, K. L., & Picton, T. W. (2007). Aging in binaural hearing begins in mid-life: Evidence from cortical auditory-evoked responses to changes in interaural phase. *Journal of Neuroscience, 27*(42), 11172–11178.

Schneider, B. A., & Pichora-Fuller, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. *In F. I. M. Craik & T. A. Salthouse*

(Eds.), Handbook of aging and cognition (2nd ed., pp. 155–220). Mahwah, NJ: Erlbaum.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123*(12), 2400–2406.

Shtyrov, Y., & Pulvermüller, F. (2002). Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport, 13*(4), 521–525.

Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R. J., & Näätänen, R. (2000). Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *Neuroimage, 12*(6), 657–663.

Sorokin, A., Alku, P., & Kujala, T. (2010). Change and novelty detection in speech and non-speech sound streams. *Brain Research, 1327*, 77–86.

Takegata, R., Tervaniemi, M., Alku, P., Ylinen, S., & Näätänen, R. (2008). Parameter-specific modulation of the mismatch negativity to duration decrement and increment: Evidence for asymmetric processes. *Clinical Neurophysiology, 119*(7), 1515–1523.

Taylor, J. K., & Burke, D. M. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. *Psychology and Aging, 17*(4), 662.

Wang, Y., Yang, X., & Liu, C. (2017). Categorical perception of Mandarin Chinese tones 1–2 and tones 1–4: Effects of aging and signal duration. *Journal of Speech, Language, and Hearing Research, 60*(12), 3667–3677.

Wang, Y., Yang, X., Zhang, H., Xu, L., Xu, C., & Liu, C. (2017). Aging effect on categorical perception of Mandarin tones 2 and 3 and thresholds of pitch contour discrimination. *American Journal of Audiology, 26*(1), 18–26.

Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage, 33*(1), 316–325.

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience, 170*(1), 223–231.

Yang, X., Wang, Y., Xu, L., Zhang, H., Xu, C., & Liu, C. (2015). Aging effect on Mandarin Chinese vowel and tone identification. *The Journal of the Acoustical Society of America, 138*(4), 411–416.

Yip, M. (2002). *Tone.* Cambridge, UK: Cambridge University Press.

Yu, K., Zhou, Y., Li, L., Su, J. A., Wang, R., & Li, P. (2017). The interaction between phonological information and pitch type at pre-attentive stage: An

ERP study of lexical tones. *Language, Cognition and Neuroscience, 32*(9), 1164–1175.

Zhao, L. (2010). *Experimental course of ERPs.* Nanjing, China: Southeast University Press.

Appendices

Appendix 1: Semantic Association Rating Questionnaire

Please rate the semantic association strength of each syllable (i.e., the degree to which meaning can be associated). Semantic associations may come from Mandarin, dialects, or foreign languages. Use the following scale: 5 = very high; 4 = relatively high; 3 = moderate; 2 = relatively low; 1 = very low. 5 is the highest score, 1 is the lowest.

[Table of syllable pairs would appear here]

Note: Because spectrally rotated nasal consonants sound extremely similar to speech, pseudo-syllables containing nasal consonants were not included in the semantic association rating.

Appendix 2: Speech Similarity Rating Questionnaire

Please rate the speech similarity of each sound (i.e., the degree to which it sounds like human speech). Use the following scale: 5 = very high; 4 = relatively high; 3 = moderate; 2 = relatively low; 1 = very low.

[Table of rotated syllable pairs would appear here]

Appendix 3: Edinburgh Handedness Inventory

Gender: Male Female

For each activity below, indicate which hand you prefer by checking the appropriate box. If the preference is very strong (i.e., you never use the other hand unless forced), check two boxes ( ). If you have no preference for a particular activity, check both boxes. Please answer all questions. If you have no experience with the object or task, leave it blank.

[Activities list would appear here]

Scoring: Count the + signs in each column. Sum left and right hand totals for the combined score. Subtract left hand total from right hand total for the difference score. Divide difference by combined score (round to 2 decimal places if needed) and multiply by 100 for the result.

Result: Below -40 = left-handed; -40 to 40 = mixed-handed; Above 40 = right-handed

Appendix 4: Mini-Mental State Examination (MMSE)

[MMSE items would appear here]

Note: Normal values should be determined by education level: Illiterate >17; Primary school >20; Secondary school and above >24.

Total Score: ___________________

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.