



Effects of Aging and Age-Related Hearing Loss on Talker Discrimination

Min Xu¹, Jing Shao², Lan Wang¹

¹Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

²Department of English Language and Literature, Hong Kong Baptist University, Hong Kong SAR, China

xumin@siat.ac.cn, jingshao@hkbu.edu.hk, lan.wang@siat.ac.cn

Abstract

Paralinguistic information is as important as linguistic information. Being familiar with talker's voice may facilitate speech perception, especially in challenging conditions. Previous studies have suggested that aging and age-related hearing loss lead to the deterioration of the phonetic and phonological processing ability. The current study aims to explore whether these two factors exert effects on the talker's voice discrimination. Three groups of participants, including young adults (YA) and older adults (OA) with and without hearing loss, were tested on talker discrimination in four types of stimuli varying in language familiarity: Mandarin real words, pseudowords, Arabic words and reversed Mandarin words. The results showed that OA with and without hearing loss performed worse than YA in both nonnative and native conditions. OA with hearing loss further performed worse than OA with normal hearing in Mandarin real word condition. These findings indicated that aging and hearing loss affected both low-level phonetic and high-level phonological processing, but hearing loss had extra effect on phonological processing. Altogether, these results implied that OA could not utilize phonetic and phonological cues as effectively as YA, and OA with hearing loss encountered more difficulties in utilizing phonological cues in talker discrimination.

Index Terms: aging, age-related hearing loss, language familiarity, talker discrimination

1. Introduction

Paralinguistic information/language independent information, such as speakers' gender, age and emotion, expressed concurrently with linguistic features/language specific information, plays an equally remarkable role in speech perception. Successful speech perception requires listeners not only to know "what is said", but also to gain additional information about the individual "who is saying".

There are interactions between talker recognition and speech understanding. On the one hand, listeners process talker information instead of normalizing it during speech comprehension [1, 2], and familiar talkers' speech is more comprehensible to listeners. On the other hand, a multitude of behavioral studies have indicated that linguistic information exerts an influence on talker recognition [3, 4, 5, 6, 7, 8]. Listeners tend to show better performance when identifying talkers speaking native language compared with foreign language. This special property of talker recognition is called language familiarity effect. For example, an early study showed that monolingual English listeners identified talkers who are speaking foreign Spanish less accurately than those who are speaking native English [8]. Another study, which included monolingual English and German listeners, also

showed similar findings [3]. The language familiarity effect was possibly related with the potential phonological processing abilities [9, 10]. This assumption has been confirmed by studies referring to dyslexics, who have poorer phonological processing abilities. It is found that dyslexics performed worse than typical controls in native language condition (English), but the group difference was absent in the unfamiliar language condition (Chinese), demonstrating that the ability of recognizing voice is associated with the ability in the perception of phonetic features and the access to phonological representations [11]. Taken together, it has been confirmed that two main factors may influence talker identification, i.e., phonological cues available in speech signal and the listeners' phonological processing abilities.

Although there have been several studies which examined the talker recognition in infants [4,12], children [13, 14] and adults [3, 5], talker discrimination in older populations was less explored. OA were proven to have difficulties in speech comprehension due to their poorer phonetic and phonological processing abilities, which also played important roles in talker recognition. To date, whether aging affects talker discrimination and whether it affects low-level phonetic and high-level phonological processing of talker's voice to the same extent is still controversial. For example, a study found that there were no significant age-related changes when OA were required to judge the sex of the speaker but ignore what was said [15]. However, another study showed that the talker identification accuracy in OA decreased when they simultaneously listened to a word as target and another meaningful word as a distractor, or in the conditions which contained more voice variability. In contrast, YA's performance was not affected, demonstrating high scores in all conditions. These findings suggested that OA may encounter difficulties not only in low-level phonetic processing but also in high-level phonological processing, especially when the task was more complex and demanding [16]. In a similar vein, another study which systematically manipulated the phonetic information, i.e., spectral and temporal cues of the talkers' vowel tokens, found that OA with normal hearing identified talkers' gender less accurately than YA. Moreover, they benefited less from the increased temporal envelope information, which suggested that in addition to the difficulties in processing phonetic information in talker's gender identification, OA benefited less from the additional phonetic cues in the demanding conditions [17].

Sensory hearing loss is caused by the peripheral encoding degradation, which may directly affect low-level auditory processing. In addition, the impoverished auditory experience may cause a deterioration of high-level phonological processing. These assumptions have been confirmed by several studies which found that hearing-impaired group performed slower and less accurate in a rhyme-judgement task

which mainly depend on phonological ability [18]. However, whether these phonetic and phonological deficits caused by sensory hearing loss affect talker discrimination received very little attention. To our knowledge, there was only one study found that three groups of listeners, including OA with and without hearing loss, as well as YA with hearing loss, performed worse than YA with normal hearing in talker identification. The follow-up regression analyses indicated that both age and hearing loss may affect the talker identification performance [19], but this study did not include a condition that can test the language familiarity effect in OA.

As mentioned above, language familiarity effect existed in talker processing, but it still lacks evidence to reveal whether aging and age-related hearing loss affect talker processing and whether these two factors affect the low-level phonetic and high-level phonological processing to the same degree. To this end, we conducted a talker discrimination experiment with four types of stimuli varying in the amount of the available phonological cues: Mandarin real words which contained both phonological and semantic information, Mandarin pseudowords which contained valid phonological information but were meaningless, Arabic words which contained unfamiliar phonological information and reversed Mandarin words which have eliminated phonological information. YA, OA with and without hearing loss were recruited. We predict YA may show language familiarity effect, performing better in the native language conditions where contain language-specific information (Mandarin real word and pseudoword conditions) than in nonnative conditions where only contain language-independent information (Arabic and reversed conditions). In addition, we predict aging has adverse influences on both low-level phonetic and high-level phonological processing in talker discrimination, which means that OA with normal hearing would perform worse than YA in both nonnative and nonnative conditions, and they would perform much worse than YA in native language condition if aging has more effects on phonological processing. As for the influence of hearing loss, we predict it exerts extra influences on phonetic and phonological processing, which means OA with hearing loss performed worse than their peers with normal hearing in all conditions. Likewise, we predict, the performance of OA with hearing loss are worse than their peers in native conditions if hearing loss has extra effects on phonological processing.

2. Method

2.1. Participants

Three groups of listeners, OA with and without hearing loss and YA participated in this experiment. All participants were from north of China and spoke standard Mandarin. None of them had any knowledge of Arabic. The hearing condition was defined by pure-tone detection thresholds at octave frequencies from 250Hz to 8000Hz, bilaterally. The group of OA with hearing loss contained 18 participants (9 females and 9 males), ranging from 57-75 years (Mean=64.47, SD=4.52). Most of them have moderate hearing loss at 2000Hz and have moderate or severe hearing loss at 4000Hz and 8000Hz. The group of OA with normal hearing contained 20 participants (13 females and 7 males), ranging from 59-70 (Mean=63.20, SD=3.41). The binaural pure-tone detection thresholds were better than 25dB HL below 4000Hz. The group of YA contained 17 participants (9 females and 8 males), ranging

from 20-27 (Mean=20.88, SD=1.80). Their binaural pure-tone detection thresholds were all below 20dB HL at all octave frequencies. The audiogram for the three groups is shown in Figure 1. All participants were paid for their participation. The experimental procedures were approved by the Human Subjects Ethics Committee of the Shenzhen Institutes of Advanced Technology, Chinese Academy of Science. Informed written consent was obtained from participants in compliance with the experiment protocols.

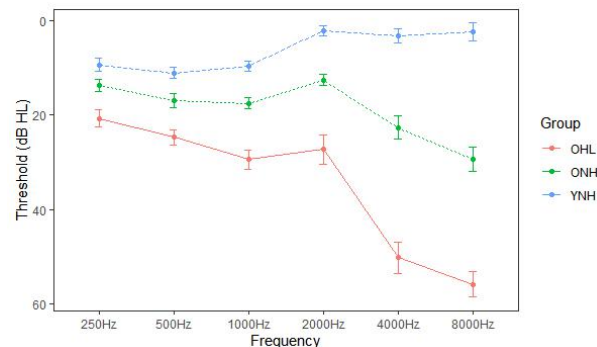


Figure 1: The audiogram of participants .

2.2. Stimuli

Four types of words (Mandarin real words, Mandarin pseudowords, Arabic real words, and reversed Mandarin words) were prepared. The Mandarin real words were commonly-used disyllabic words. The Mandarin pseudowords were also disyllabic. The syllables and tones both exist in Mandarin, but the combinations of the syllables and tones are illegitimate and therefore meaningless (e.g., /le35 nuo214/). Likewise, disyllabic words were chosen as Arabic stimuli. The reversed Mandarin words were transformed Mandarin real words. Each type of word set consisted of 120 words. The speakers were six female Mandarin-L1/Arabic-L2 bilinguals, and their Arabic proficiency is close to native speakers. They produced the first three types of words (Mandarin real words and pseudowords, and Arabic real words).

The talkers were given the word lists and enough time to get familiar with the pronunciation of each word, especially the Mandarin pseudowords. Three types of stimuli were recorded separately, and the words were presented randomly to talkers through E-prime 2.0. When they recorded the Chinese real words, Chinese characters and pinyin were presented simultaneously, but when came to the Chinese pseudowords condition, only pinyin were presented. Arabic scripts were presented when they recorded. All the words were displayed on the middle of the computer screen one by one. The talkers were asked to read the words loudly and naturally and repeat each words three times. The talkers were also instructed to pause between every two words in order to facilitate later segmentation. The sampling rate of the recordings was 22050Hz with 16 bits per sample. The whole recording contained 6480 words in total (3 word types \times 120 words \times 6 talkers \times 3 repetitions).

The clearest word pronounced by each talker was chosen. All the later manipulation was done with Praat. The target words were first segmented and then their duration and intensity were normalized. The duration was normalized separately based on the mean duration in four conditions using PSOLA, which can preserve other acoustic features, like fundamental frequency, formant and amplitude. The whole

Mandarin real words were normalized to 712ms, the Mandarin pseudowords to 875ms, and the Arabic words to 732ms. And then all four types of words were normalized to 75dB. The Reversed words were created by time reversing the Mandarin real words, so the reversed words had the same duration and intensity as the Mandarin real words.

2.3. Procedure

The stimuli presentation and response collection were through psychology software tools: E-prime 2.0. The task was a two-alternative forced choice talker discrimination. Each type of stimuli corresponded to a block of discrimination task, resulting in four blocks in total. Each block contained 60 talker pairs, 30 same talker pairs and 30 different talker pairs. The 30 different talker pairs consisted of two groups of 15-pairs. The first group were 15 non-repeating pairs formed by pairing six different talkers. The second group were formed through reversing the pair sequence of the first group (e.g., talker1-talker2, talker2-talker1). Six same talker pairs repeated 5 times formed the 30 same talker pairs. In order to increase the task difficulties and encourage listeners to judge the talker information in a more complex way, the two words in each trail, both in same pairs and different pairs, were always different which maximized the differences of the acoustic and phonological cues.

Auditory stimuli were presented via Sennheiser 280 headphones binaurally to listeners. The stimuli were presented to each participant in 75dB SPL. In each trail, listeners first saw a fixation on the screen for 500ms, and then listened to the two disyllable words with a 500ms inter-stimulus interval. The listeners were instructed to press “left arrow” or “right arrow” to judge whether the two words were spoken by the same talker or different talkers. The time limit for the response was 5 seconds, the next trail occurred automatically if no responses were detected. To prevent fatigue of listeners, within-condition, for every 20 trails, and between-condition breaks were given. Four blocks were presented to listeners in a counterbalanced way and these sequences were matched one-to-one among three groups. And 60 pairs in each block were presented in a random sequence. Before the formal experiment, the participants were provided a practice session to familiarize themselves with the experimental procedures and the stimuli types. The practice session also included four blocks consisted of different types of stimuli, and each blocks included four trails. No feedback was given to participants either in practice session and form session.

2.4. Data analysis

The data was analyzed in terms of overall discrimination performance and reaction time (RT). The discrimination performance for each listener was based on d' rather than on the percentage correct. The calculation of d' was the z-score value of hits minus the z-score value of false alarms. Hits were responses of “different” to different talker pairs, and false alarms were responses of “different” to same talker pairs. We constructed a Linear mixed-effected models of d' scores with group (YA, OA with and without hearing loss), condition (Mandarin real word and pseudoword, Arabic word, and reversed word) and the two-way interaction as fixed effects and with by-subject as random effects. RT was the time between the end of the final stimulus and the response was made. Another Linear mixed-effected models were fitted on the log-transformed RT with group and condition and the two-way interaction as fixed effects and with by-subject and each

pair of stimulus as random effects. All the analysis were performed with R [20], and the lme4 package [21], the lmer package [22], and the lsmeans package [23] were used.

3. Results

For the d' scores, there were significant main effects of group ($F(2,44)=23.766$, $p<.001$) and condition ($F(3,132)=12.809$, $p<.001$). The two-way interaction effect was approaching significant ($F(6,132)=1.862$, $p=0.092$). The detailed d' scores are shown in Figure 2. Pairwise comparisons showed in the Mandarin condition, the d' scores were highest in YA, followed by OA with NH and OA with HL. The differences among the three groups were significant ($ps<.001$). In the pseudoword, Arabic and reversed conditions, the d' scores of the two OA groups were significantly lower than the young group ($ps<.001$), while the d' scores between the OA groups were not significant. As for the condition effect within each group, it showed that for YA, the d' scores in Mandarin were significantly higher than in reversed and marginally significant higher than in pseudoword condition. Besides, the d' scores in Arabic condition were significant higher than in reversed condition and in pseudoword condition. The performance of OA with normal hearing were similar with YA, excepted that the d' scores in Arabic condition were not significantly higher than in pseudoword condition. However, for OA with hearing loss, there were no significant differences in four conditions.

For the RT, there were a main effect of condition ($F(3,436.7)=6.132$, $p<.001$) and the two-way interaction between group and condition ($F(6,12970.9)=4.627$, $p<.001$). Pairwise comparisons showed that there were no significant differences among three groups in four conditions. However, we found that for YA, RTs elicited in the Mandarin condition were significantly shorter than in the pseudoword and reversed condition ($ps<.05$). The RTs elicited in the pseudoword condition were significantly longer than in the Arabic and the Reversed condition, while the difference between the Arabic and the reversed condition was not significant. For the OA with normal hearing, there were different patterns compared with the YA. RTs elicited in the Mandarin and pseudoword condition were significantly longer than in the reversed condition ($ps<.05$). No other effects were significant. For the OA with hearing loss, no effects were found significant.

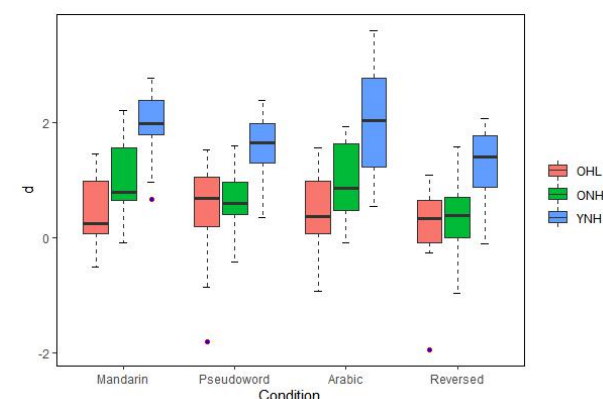


Figure 2: Talker discrimination accuracy.

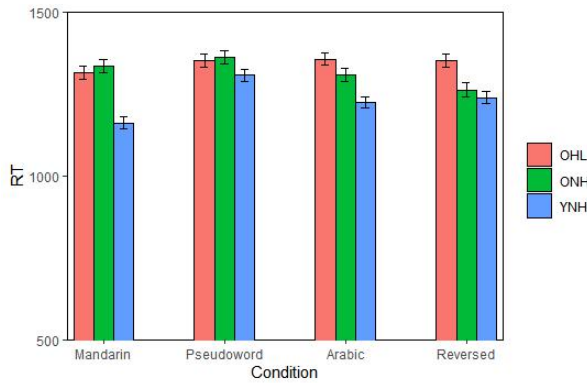


Figure 3: Talker discrimination RT.

4. Discussion

We predict that YA could perform better in native language conditions (Mandarin real word and pseudoword conditions) than nonnative conditions (Arabic and reversed conditions). We also predict that OA with normal hearing perform worse than YA in native and nonnative conditions and they perform much worse than YA in native conditions if aging has more effects on phonological processing. Furthermore, we predict OA with hearing loss perform worse than OA with normal hearing in all conditions. Similarly, OA with hearing loss may perform much worse than their peers in native language. Our findings confirmed parts of the above predictions. For YA, Mandarin real word condition elicited the most highest d' scores and most shortest RT, and reversed condition elicited the lowest d' scores. OA with and without hearing loss got lower scores than YA in all four conditions. However, what was unexpected was that YA's d' scores and RT in Arabic condition were similar to those in Mandarin real word condition, and they employed longer time in pseudoword condition. Lastly, OA with hearing loss were lower than their peers only in Mandarin real word condition.

Previous studies found the language familiarity plays an important role in talker identification [3, 6, 8, 15, 18]. Our findings in YA confirmed that the language familiarity effect also exist in talker discrimination task which primarily depended on the comparison of the phonetic information between the two stimuli within one trial. It indicated that YA listeners employed top-down phonological processing to some extent even in an undemanding task. Contrary to our expectation, Arabic condition elicited comparably high accuracy as Mandarin real word condition and pseudoword condition elicited the longest RT. The similar performance between Arabic and Mandarin real word conditions were in line with a previous study where Chinese typical listeners performed equally in Mandarin real word and Arabic conditions, and pseudoword condition also elicited longer RT [24]. Another study got similar results that native English listeners performed almost the same for other language trials (German, Finnish and Mandarin) compared with English trials, suggesting that the talker discrimination could be across language [25]. Although the accuracy was comparable, YA spent more time on processing unfamiliar phonological information in Arabic condition, which can be considered a sign of language familiarity effect. The possible reason for the longer RT in pseudoword condition is that despite the pseudowords contained real phonological components in Mandarin such as consonants, vowels and lexical tones, the

combinations of these elements were meaningless and illegitimate, conflicting with listeners' existing phonological knowledge and resulting in a longer time to process the phonological information.

The finding that OA with normal hearing performed worse than YA in all conditions suggested that although older adults have normal hearing, their abilities to compare low-level phonetic cues as well as utilize high-level phonological cues are still impaired to some extent. It was consistent with the previous findings that OA discriminated talkers' gender worse than YA [15] and the talker identifications of OA were decreased in a more competing environment where phonetic and phonological interference increased [18]. These deficits in talker discrimination also occurred in other linguistic processing, such like categorical perception of speech [26] and vowel/consonant identification [27]. Furthermore, OA demonstrated a weak language familiarity effect, but they spent more time in Mandarin real word and pseudoword condition than in reversed condition compared with YA, indicating OA with normal hearing could still access and utilize phonological cues to some extent to facilitate the talker discrimination, but their processing speed declined. There was a transmission deficit hypothesis proposing that the declined phonological abilities in OA was attributed to the reduced connection with previous phonological knowledge [28]. In the current study it is possible that OA may need more time in native conditions to active existing phonological memory and build the connection between the incoming stimuli and the stored phonological representations.

OA with hearing loss could not discriminate different talkers well when compared to YA. Furthermore, they performed worse than their peers in Mandarin real word condition, indicating that independent from aging, age-related hearing loss leads to severe declines in utilizing phonological information to facilitate talker discrimination. The finding was consistent with a study which showed that hearing loss was related to the semantic long term memory which were based on phonological ability [29]. All of these findings confirmed that hearing loss does not simply result in poor sensory sensitivity, it also affects high-level phonological abilities. The poor hearing condition lead to insufficient input for OA, which may further affect their high-level, such as phonological, semantic processing and slow down their top-down processing speed.

5. Conclusions

In the current study, YA demonstrated language familiarity effect, performing better in native conditions than nonnative conditions. Furthermore, both aging and hearing loss contributed to the talker discrimination but did so differently. They affected both the low-level phonetic processing and the high-level phonological processing, however, hearing loss exerted an additional influence on phonological processing in talker discrimination. Our findings further expanded the understanding of aging and age-related hearing loss.

6. Acknowledgements

This work was supported by grant from the National Natural Science Foundation of China (NSFC: 11904381). We thank Dr. Zhang Caicai for her constructive comments on the stimulus design.

7. References

- [1] S. D. Goldinger, "Words and voices: Episodic traces in spoken word identification and recognition memory," *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 22, no. 5, pp. 1166–1183, 1996.
- [2] K. Johnson, "The role of perceived speaker identity in F 0 normalization of vowels," *The Journal of the Acoustical Society of America*, vol. 88, no. 2, pp. 642–654, 1990.
- [3] J. P. Goggin, C. Thompson, G. Strube, and L. R. Simental, "The role of language familiarity in voice identification," *Memory & cognition*, vol. 19, no. 5, pp. 448–458, 1991.
- [4] E. K. Johnson, E. Westrek, T. Nazzi, and A. Cutler, "Infant ability to tell voices apart rests on language experience," *Developmental Science*, vol. 14, no. 5, pp. 1002–1011, 2011.
- [5] S. J. Winters, S. V. Levi, and D. B. Pisoni, "Identification and discrimination of bilingual talkers across languages," *The Journal of the Acoustical Society of America*, vol. 123, no. 6, pp. 4524–4538, 2008.
- [6] A. G. Goldstein, P. Knight, K. Bailis, and J. Conover, "Recognition memory for accented and unaccented voices," *Bulletin of the Psychonomic Society*, vol. 17, no. 5, pp. 217–220, 1981.
- [7] T. K. Perrachione and P. C. M. Wong, "Learning to recognize speakers of a non-native language: Implications for the functional organization of human auditory cortex," *Neuropsychologia*, vol. 45, no. 8, pp. 1899–1910, 2007.
- [8] C. P. Thompson, "A language effect in voice identification," *Applied Cognitive Psychology*, vol. 1, no. 2, pp. 121–131, 1987.
- [9] D. Fleming, B. L. Giordano, R. Caldara, and P. Belin, "A language-familiarity effect for speaker discrimination without comprehension," *Proceedings of the National Academy of Sciences*, vol. 111, no. 38, pp. 13795–13798, 2014.
- [10] E. K. Johnson, L. Bruggeman, and A. Cutler, "Abstraction and the (Misnamed) Language Familiarity Effect," *Cognitive Science*, vol. 42, no. 2, pp. 633–645, 2018.
- [11] T. K. Perrachione, S. N. Del Tufo, and J. D. E. Gabrieli, "Human voice recognition depends on language ability," *Science*, vol. 333, no. 6042, p. 595–595, 2011.
- [12] N. Fecher and E. K. Johnson, "Effects of language experience and task demands on talker recognition by children and adults," *The Journal of the Acoustical Society of America*, vol. 143, no. 4, pp. 2409–2418, 2018.
- [13] S. V. Levi and R. G. Schwartz, "The Development of Language-Specific and Language-Independent Talker Processing," *Journal of Speech, Language, and Hearing Research*, vol. 56, no. 3, pp. 913–920, 2013.
- [14] N. S. Dailey, E. Plante, and R. Vance, "Talker discrimination in preschool children with and without specific language impairment," *Journal of Communication Disorders*, vol. 46, no. 4, pp. 330–337, 2013.
- [15] S. Jerger, F. Pirozzolo, J. Jerger, R. Elizondo, S. Desai, E. Wright, and R. Reynosa, "Developmental trends in the interaction between auditory and linguistic processing," *Perception & Psychophysics*, vol. 54, no. 3, pp. 310–320, 1993.
- [16] J. Rossi-Katz and K. H. Arehart, "Message and talker identification in older adults: Effects of task, distinctiveness of the talkers' voices, and meaningfulness of the competing message," *Journal of Speech, Language, and Hearing Research*, vol. 52, no. 2, pp. 435–453, 2014.
- [17] K. C. Schwartz and M. Chatterjee, "Gender identification in younger and older adults: use of spectral and temporal cues in noise-vocoded speech," *Ear and hearing*, vol. 33, no. 3, pp. 411–420, 2012.
- [18] U. Andersson and B. Lyxell, "Working memory deficit in children with mathematical difficulties: A general or specific deficit?," *Journal of experimental child psychology*, vol. 96, no. 3, pp. 197–228, 2007.
- [19] V. Best, J. B. Ahlstrom, C. R. Mason, E. Roverud, T. K. Perrachione, Jr, G. Kidd and J. R. Dubno, "Talker identification: Effects of masking, hearing loss, and age," *Journal of the Acoustical Society of America*, vol. 143, no. 2, pp. 1085–1092, 2018.
- [20] R Core Team, "R: A language and environment for statistical computing," 2013.
- [21] D. Bates, M. Maechler, and B. Bolker, "lme4: Linear mixed-effects models using Eigen and R interfaces." R package version 0.999375-42. 2011. R Project.
- [22] A. Zeileis and T. Hothorn, "Diagnostic Checking in Regression Relationships," *R News*, vol. 2, no. 3, pp. 7–10, 2002.
- [23] R. V. Lenth, "Least-squares means: The R package lsmeans" *Journal of Statistical Software*, vol. 69, no. 1, pp. 1–33, 2016.
- [24] J. Shao, L. Wang, and C. Zhang, "Talker processing in Mandarin-speaking congenital amusics," *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 5, pp. 1361–1375, 2020.
- [25] M. Wester, "Talker discrimination across languages," *Speech Communication*, vol. 54, no. 6, pp. 781–790, 2012.
- [26] S. L. Mattys and O. Scharenborg, "Phoneme categorization and discrimination in younger and older adults: A comparative analysis of perceptual, lexical, and attentional factors," *Psychology and Aging*, vol. 29, no. 1, pp. 150–162, 2014.
- [27] S. Gordon-Salant, G. H. Yeni-Komshian, P. J. Fitzgibbons, and J. Barrett, "Age-related differences in identification and discrimination of temporal cues in speech segments," *The Journal of the Acoustical Society of America*, vol. 119, no. 4, pp. 2455–2466, 2006.
- [28] D. G. MacKay and L. E. James, "Sequencing, Speech Production, and Selective Effects of Aging on Phonological and Morphological Speech Errors," *Psychology and Aging*, vol. 19, no. 1, pp. 93–107, 2004.
- [29] J. Rönnerberg, H. Danielsson, M. Rudner, S. Arlinger, S. Sternäng, A. Wahlin, and L. G., Nilsson, "Hearing loss is negatively related to episodic and semantic long-term memory but not to short-term memory," *Journal of Speech, Language, and Hearing Research*, vol. 54, no. 2, pp. 705–726, 2011.