INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Effects of hearing loss and amplification on Mandarin consonant perception

Huali Zhou[1], *Xianming Bei*[2], *Huanshi Xu*[3], *Yingfeng Sun*[3], *Zhixin Lin*[3],
Nengheng Zheng[1], *Qinglin Meng*[4]

[1]Guangdong Key Lab of Intelligent Information Processing, CEIE, Shenzhen University, China
[2]Faulty of Chinese Language and Culture, Guangdong University of Foreign Studies, China
[3]Bailing Hearing Clinic, Guangzhou, 510641, Guangdong, China
[4]School of Physics and Optoelectronics, South China University of Technology, China

`nhzheng@szu.edu.cn, mengqinglin@scut.edu.cn`

Abstract

This study investigates the effects of hearing loss and amplification on Mandarin consonant perception. 44 listeners with varying degrees of hearing loss were tested, both with and without the use of hearing aids. Consonant recognition was strongly correlated with the hearing threshold ($r = -0.87$), and was significantly improved by hearing-aid amplification (by more than 20% in group means) but was still not perfect. The underlying reasons are discussed. Furthermore, confusion patterns were analyzed and compared with those of normal-hearing listeners in the literature. The most challenging Mandarin consonants for hearing-impaired listeners include consonants with a spectral center of gravity in the high-frequency range (such as **s** and **z**), consonants with short duration (such as **b**, **d**, and **g**), and aspirated stops (such as **p**, **t**, and **k**). The findings of this study contribute to a better understanding of the difficulties experienced by Mandarin-speaking listeners with hearing loss.

Index Terms: speech recognition, hearing aid, confusion pattern, hearing loss

1. Introduction

Vowels and consonants both carry important information that is crucial for determining the meaning of spoken language and enabling effective communication. However, hearing loss can significantly affect their perception, with consonants being more vulnerable to impairment [1], leading to difficulties in speech perception, which is the main complaint of listeners with hearing loss. Hearing aid amplification can improve speech perception by restoring the audibility of previously imperceptible acoustic cues and compressing the range of sound intensity into an appropriate one. Numerous studies have investigated the impact of hearing loss and amplifications on the perception of English consonants [2, 3, 4, 5].

Mandarin, one of the most widely spoken languages in the world, has a unique set of consonants that differ phonetically from English consonants in terms of articulation manner and place. Based on the articulation manner, the contrast between voiced and voiceless in stops and affricates in English does not exist in Mandarin. Instead, the contrast is between aspirated and unaspirated stops and affricates in Mandarin. Based on the articulation place, Mandarin has alveolo-palatal consonants (j /tɕ/, q /tɕʰ/, x /ɕ/) and retroflex consonants (zh /tʂ/, ch /tsʰ/, sh /ʂ/, r /ʐ/) that do not exist in English. On the other hand, English has postalveolar consonants (/ʤ, ʧ, ʃ, ʒ/) that are not found in Mandarin. Variations in the manner and place of articulation can result in distinct acoustic information and may be differently affected by hearing loss. For example, Mandarin has three-way contrasts among sibilant fricatives (s /s/, x /ɕ/, sh /ʂ/ and affricates like z /ts/, j /tɕ/, zh /tʂ/ and c /tsʰ, q /tɕʰ/, ch /tʂʰ/. For people with high-frequency hearing loss, the three-way contrasts in high-frequency sounds may pose more challenges than the two-way contrast of sibilant fricatives found in English (such as /s/ and /ʃ/).

Despite the large Mandarin-speaking population, data on Mandarin consonant perception by listeners with hearing loss are surprisingly limited. The confusion patterns of Mandarin consonants among normal-hearing listeners have been well-established in previous research [6, 7]. Although there have been some reports on the performance of Mandarin consonant perception among listeners with hearing loss in studies involving children [8, 9] or evaluating hearing aid algorithms [10, 11, 12, 13], a comprehensive analysis of confusion patterns is lacking.

This study aims to fill the above gap in the literature by exploring the impact of hearing loss and hearing-aid amplification on the perception of Mandarin consonants and to compare the results with existing findings on English consonants. We present Mandarin consonant recognition data of individuals with varying degrees of hearing loss, both with and without the use of hearing aids. Furthermore, confusion patterns are analyzed and compared to those of normal-listening listeners.

2. Methods

2.1. Speech materials

The stimulus set included 21 **/Cā/** syllables (**/Ciā/** for **j**, **q**, and **x**) constructed from the exhaustive combinations of 21 Mandarin initial **C**onsonants (**b**, **p**, **m**, **f**, **d**, **t**, **l**, **g**, **k**, **h**, **j**, **q**, **x**, **zh**, **ch**, **sh**, **z**, **c**, **s**, **y**, **w**) with final vowel **a** in Tone 1 (hereinafter, Mandarin consonants and vowels are shown in boldface). In Chinese phonetics research, the sounds **y** and **w** are generally not classified as initials, but in actual pronunciation, there is often a consonantal component and auditory characteristic at the beginning of the syllable where **y** and **w** are located. These sounds are actually pronounced as approximants and are also examined **y** and **w** in this study. The consonants **n** and **r** were not included because their combination syllables with **a** in Tone 1 do not exist or are not commonly used in conversational speech. Materials were recorded in an anechoic room in a natural manner by a female speaker with professional training in broadcasting and stored as 16-bit .wav files with a sampling rate of 44.1 kHz.

2.2. Participants

Forty-four hearing-impaired listeners (23 females) with sensorineural hearing loss participated in the experiment. They were all native Mandarin speakers aged 9 to 80 years with a mean age of 33.7±17.8 years. Pure-tone audiometry was performed for all listeners and the results are shown in Fig. 1.

10.21437/Interspeech.2023-536

The group means the 4-frequency pure-tone average threshold (4fPTA) at octave frequencies from 500 to 4000 Hz [14] was 63 ± 21 and 56 ± 18 dB HL for the left and right ear, respectively. All except one listener had bilateral hearing loss. Based on better-ear 4fPTA, one listener had normal hearing (as defined by the better-ear 4fPTA of < 20 dB HL), seven had mild hearing loss (20-34 dB HL), 26 had moderate hearing loss (35–64 dB HL), and ten had severe hearing loss (65–95 dB HL).

Among the 44 listeners, 39 had used hearing aids before participating in the study. All listeners were tested without hearing aids (the unaided condition) and 38 of them were tested with hearing aids (the aided condition). In the aided condition, listeners who were using hearing aids were tested using their own hearing aids (price range from 50 to 6000 USD) with their daily used parameter settings. For listeners who were not using hearing aids, they were fitted with a pair of hearing aids in both ears for the experiment. Parameters of the hearing aids were prescribed by the fitting software based on the listeners' audiograms.

2.3. Procedure

Each listener was tested under two conditions: unaided and aided. The order of conditions was randomized across listeners. The experiment was carried out in a soundproof room. Stimuli were presented through a loudspeaker located 1 m in front of listeners, at a presentation level of about 65 dB SPL. The tests were administered through a customized computer program. A graphical user interface (GUI) was built to present the stimuli and collect listeners' responses in a 21-alternative forced-choice (21AFC) paradigm. In each trial, a syllable was presented and 21 buttons labeled with the Chinese characters as well as the pinyin were shown on the GUI. The listeners were instructed to select the perceived syllable from the 21 buttons using a computer mouse. The presentation order was randomized. Before the formal test, a brief training was conducted to familiarize listeners with the procedure. No feedback about the correctness of the response was given to listeners. The rate of correct responses was calculated as the test result in each condition and the confusion pattern was analyzed.

2.4. Confusion pattern analysis

The similarity (Eq. 1) between two consonants was used to quantify the difference between confusion patterns of the unaided and aided conditions.

$$S_{ij} = \frac{p_{ij} + p_{ji}}{p_{ii} + p_{jj}} \tag{1}$$

where p_{ij}, p_{ji}, p_{ii}, and p_{jj} denote the rate of consonant i perceived as j, j perceived as i, i correctly perceived, and j correctly perceived, respectively. S_{ij} denotes the similarity between consonant i and j with a value ranging from 0 to 1. A higher value of S_{ij} indicates that the two consonants are more similar, and thus it is more difficult to discriminate between them. For any consonant i and j, we have $S_{ij} = S_{ji}$, and $S_{ii} = S_{jj} = 1$. Therefore, the similarity matrix S is a symmetric matrix with diagonal elements equal to 1. Hierarchical clustering based on the average similarities was used to cluster the consonants based on their similarities using the *hclust()* function with a parameter *method* = *"average"* in R.

3. Results

3.1. Recognition rates

Fig. 2(a) shows unaided data from all listeners ($n = 44$). Six listeners did not take the aided test, and their results were excluded from further analysis. Fig. 2(b)-(e) show data from listeners who finished testings in both unaided and aided conditions ($n = 38$).

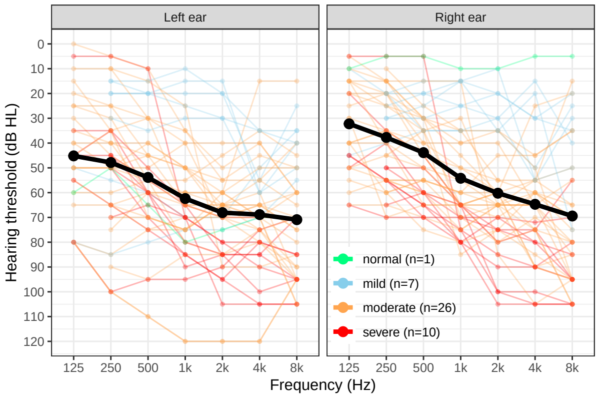

Figure 1: *Audiograms of the 44 listeners. Each thin line represents one listener, and the thick black lines show group means. The line colors indicate the degree of hearing loss based on better-ear 4fPTAs.*

3.1.1. *Unaided performance*

Consonant recognition was significantly impacted by hearing loss. Fig. 2(a) depicts a strong correlation between recognition rate and hearing threshold based on better-ear 4fPTA ($r = -0.87, p < 0.001, R^2 = 0.76$, Pearson's correlation). The mean recognition rates were 100%, 85%, 38.9%, and 7.7%, for listeners with normal hearing, and mild, moderate, and severe hearing loss, respectively (the inserted graph in Fig 2(a)). The regression line $y = -1.584x + 123.6$ predicted a 15% decrease in recognition rate for every 10 dB increase in the better-ear 4fPTA, and a listener with a better-ear 4fPTA of 40 dB HL has a predicted consonant recognition score of 50%.

3.1.2. *Unaided vs. aided performance*

As shown in Fig. 2(b), hearing-aid amplification produced a significant improvement of 22 percentage points ($t(37) = -5.941, p < 0.001$, paired-sample t-test). The group mean recognition rate increased from 41 ± 30 (standard deviation, SD) % in the unaided condition, to 63 ± 20 % in the aided condition. On the individual level, 32 out of 38 listeners had a higher score in the aided condition than in the unaided condition. The only 6 listeners who did not benefit from amplification were mostly listeners with mild hearing loss who did not use hearing aids in their daily life.

Fig. 2(c) show the effect of amplification across different degrees of hearing loss. Amplification substantially improved the recognition rates for listeners with moderate and severe hearing loss (by 26%, $t(23) = 6.57, p < 0.0001$ and 32%, $t(6) = 2.98, p = 0.02$, respectively, paired-sample t-test). However, for listeners with mild hearing loss, no significant improvement was observed (unaided 85% vs. aided 83%, $t(6) = -1.86, p = 0.642$).

3.1.3. *Recognition of individual consonants*

Fig 2(d) and (e) show the recognition rates for each individual consonant in the unaided and aided conditions, respectively. All consonants were affected to varying extents by hearing loss, and amplification provided varying levels of improvement. Meanwhile, the performance gap among the three degrees of hearing loss was also narrowed by amplification. In both conditions, **m**

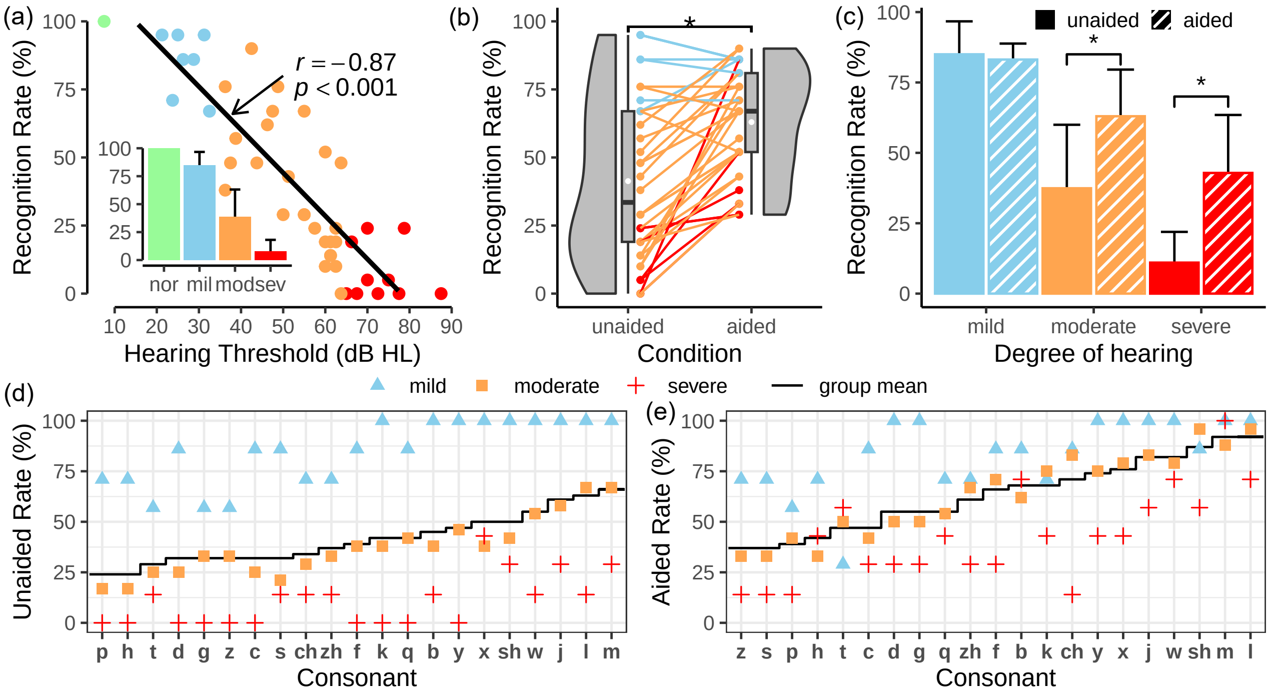

Figure 2: *Recognition rate results. (a). Scatter plot and the regression line of unaided results VS. hearing thresholds from all listeners* ($n = 44$)*. The inset shows the mean unaided results grouped by the degree of hearing loss. (b). Comparisons between unaided and aided results. Each line represents one listener. The white dots inside the boxplots show the group means. (c). Unaided and aided mean results and standard deviations grouped by the degree of hearing loss. (d). Unaided results of each individual consonant. (e). Aided results of each individual consonant. The color in all subgraphs was consistently set to indicate the degrees of hearing loss.*

and **l** had the highest recognition rates, while **p**, **h**, **t**, **d**, **g**, **z**, **c**, and **s** had the lowest recognition rates.

3.2. Confusion patterns

Results of the clustering based on similarities (see Section 2.4 for the calculation of similarities) are shown in Fig. 3(a) and (b). Similarity levels are indicated by the horizontal line segments connecting two subclusters. To show the typical confusion patterns of hearing aid users, only data from listeners who use hearing aids in their daily life were included here. Data from the seven listeners with mild hearing loss were excluded. To have a direct comparison with results by normal-hearing listeners, data (measured in several transmission conditions, including various degrees of additive noise, various sound intensities, and various frequency bandwidths) from Fig. 2-7 in reference [7] are extracted and redrawn in Fig. 3(c). Recall that we did not use the consonants **r** and **n** because their combined syllables with the vowel **a** are not commonly used in conversational speech and therefore not familiar to the listeners.

The main findings comparing the three clusters in Fig. 3 are as follows.

- First, similarities in normal-hearing listeners are much lower than in hearing-impaired listeners[1].
- Second, normal-hearing listeners are dominantly confused by the articulation place, i.e., they show mostly confusion among consonants with the same articulation manner but with different articulation places. In the main confusions shown in Fig. 3(c), **c**, **ch**, **k**, **p**, and **t** are all aspirated sounds; **f**, **h**, **s**, and **sh** are all fricatives; **g**, **z**, **zh**, **b** and **d** are all unaspirated sounds; and **m** and **n** are both nasal sounds. Confusion by the articulation manner (within the same articulation place) does exist among **j**, **q**, and **x** but with a weaker similarity level compared to those within the same articulation manner. In contrast, hearing-impaired listeners demonstrate strong confusion by the articulation manner in addition to strong confusion by the articulation place. As Fig. 3(b) shows, they are confused not only between **b** and **d**, and between **p** and **t** which differ in the articulation place, but also between **s** and **z** which differ in the articulation manner.
- Third, the highest similarity (most likely to confuse) occurred between unaspirated stops **b** and **d** for normal-hearing listeners. But for hearing-impaired listeners, it was between aspirated stops **p** and **t** when aided, or between aspirated stops **t** and fricative **h**, although the confusion between unaspirated stops **b** and **d** was still there, its level of similarity was much lower. This finding suggests that for hearing-impaired listeners, aspirated stops are more difficult to recognize compared to unaspirated stops, while normal-hearing listeners experience the opposite.

4. Discussion

This study investigates the effects of hearing loss and amplification on Mandarin consonant perception. Unaided and aided consonant recognition data from a group of Mandarin-speaking listeners with various degrees of hearing loss were presented. Furthermore, confusion patterns were analyzed based on similarities among consonants and compared with those of normal-hearing listeners reported in the literature.

---

[1] Note the different scales used in the three sub-graphs. Similarity levels reflect the difficulty of identifying consonants. The similarity level of near 1 indicates the extreme difficulty faced by hearing-impaired listeners without the use of hearing aids.

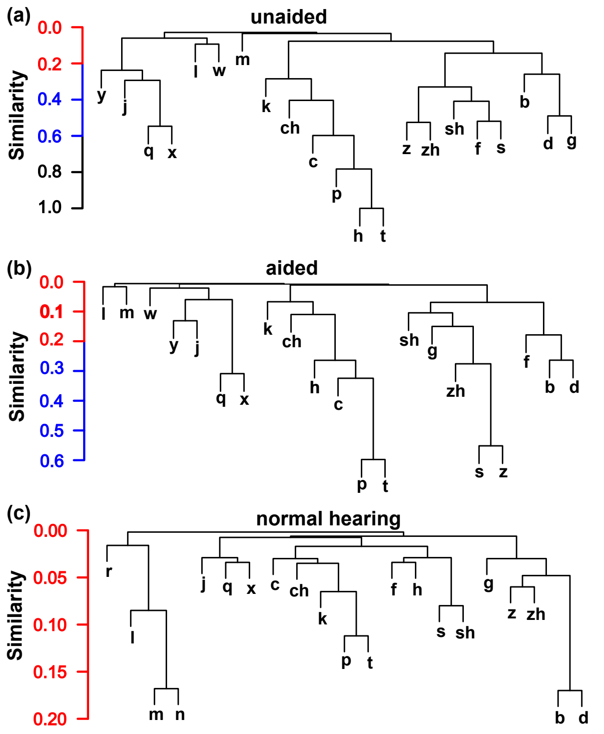

Figure 3: *Results of similarity clustering. (a) and (b) show confusion data in the unaided and aided conditions in this study, respectively. (c) shows data from normal-hearing listeners derived from Fig.2-7 in [7].*

4.1. Effects of hearing loss

The hearing threshold is strongly correlated with Mandarin consonant recognition rates ($r = -0.87$), which is consistent with reports with English consonants in the literature [4]. Listeners with hearing loss demonstrate different confusion patterns from normal-hearing listeners, which is also consistent with previous reports in studies with English-speaking listeners [2, 15, 3].

To the best of our knowledge, this is the first study to report comprehensive Mandarin consonant confusion in terms of similarity clustering from listeners with hearing loss. The acoustic characteristics of Mandarin consonants account for the results observed in this study. Voiced consonants with the spectral center of gravity located in the lower frequency range such as **m** and **l** are always well perceived even for hearing-impaired listeners. Three types of consonants are difficult for hearing-impaired listeners to recognize even with the use of hearing aids.

- consonants with the spectral center of gravity located in the high-frequency range, such as **s** and **z**. Most hearing-impaired listeners often have more severe hearing loss in the high-frequency region, making it difficult to distinguish these consonants.
- consonants with short duration, such as **b**, **d**, and **g**. The short duration and rapid pronunciation of these consonants make them unlikely to produce a deep auditory impression. These consonants are difficult to identify even for normal-hearing listeners.
- aspirated stops, such as **p**, **t**, and **k**. It is already challenging to differentiate the beginning stops segment, and the acoustic features of the following aspirating segments are similar, causing difficulties in their perception. Furthermore, listeners may easily hear **p**, **t**, and **k** as **h** if the beginning stops are not distinct and listeners only hear the following aspirating segments.

The various transmission conditions when testing normal-hearing listeners in [7, 6] mainly affect audibility. In contrast, hearing loss leads to impairments in many other aspects in addition to audibility, e.g., reduced frequency selectivity and temporal resolution [16]. These impairments contribute to the differences in confusability observed in this study and in [7, 6].

4.2. Effects of amplification

The amplification provided by hearing aids produced a significant improvement of 22% in this study, which is comparable to that observed in English consonant recognition reported in the literature [17]. There is still room for improvement in hearing aid performance, as there is still a gap in consonant perception compared to normal listeners even after using a hearing aid [18, 9, 19]. In terms of the confusion pattern, amplification substantially reduced the similarity levels among consonants and slightly changed the confusion patterns.

The potential underlying reason for the performance gap between hearing-impaired listeners with the use of hearing aids and normal-hearing listeners could be multifaceted. The first is that although hearing aids provide amplification to resolve the problem of audibility, they inevitably introduce compression because of the narrowed dynamic range of hearing-impaired listeners. The subtle acoustic differences between consonants reliably used by normal-hearing listeners were distorted and not recovered by hearing aid processing, leading to confusion in consonants with similar acoustic features. Second, the acoustic characteristics and fitting of hearing aids may also affect the speech perception of hearing-aid users. Different types of hearing aids have varying levels of technology and acoustic characteristics, and the fitting of the hearing aid can affect its effectiveness in improving speech perception. In addition, several other factors such as the listener's age and cognitive abilities, and experience with hearing aids may also play a role.

The findings of this study have important implications for the development of speech rehabilitation strategies for individuals with hearing loss, particularly in Mandarin-speaking populations. Future research should aim to further explore the underlying mechanisms of confusion and determine the most effective methods for reducing or eliminating them.

5. Conclusions

Hearing loss significantly impacts Mandarin consonant perception. Amplification can produce an improvement of more than 25% in group means for listeners with moderate-to-severe hearing loss, but not perfect performance. Hearing-impaired listeners have different confusion patterns compared to listeners with normal hearing. Specifically, they are often confused with three types of Mandarin consonants, namely consonants with a spectral center of gravity in the high-frequency range (such as **s** and **z**), consonants with short duration (such as **b**, **d**, and **g**), and aspirated stops (such as **p**, **t**, and **k**).

6. Acknowledgements

We thank all the participants who generously donated their time to participate in this study. This work was supported by the Guangdong Basic and Applied Basic Research Foundation Grant (2022A1515011361) and the Shenzhen Fundamental Research Program (20220809191805001). Nengheng Zheng and Qinglin Meng are the corresponding authors.

7. References

[1] E. Owens, C. B. Talbott, and E. D. Schubert, "Vowel discrimination of hearing-impaired listeners," *Journal of speech and hearing research*, vol. 11, no. 3, pp. 648–655, 1968.

[2] R. C. Bilger and M. D. Wang, "Consonant confusions in patients with sensorineural hearing loss," *Journal of Speech and Hearing Research*, vol. 19, no. 4, pp. 718–748, 1976.

[3] J. R. Dubno, D. D. Dirks, and L. R. Langhofer, "Evaluation of hearing-impaired listeners using a nonsense-syllable test ii. syllable recognition and consonant confusion patterns," *Journal of Speech, Language, and Hearing Research*, vol. 25, no. 1, pp. 141–148, 1982.

[4] D. L. Woods, T. Arbogast, Z. Doss, M. Younus, T. J. Herron, and E. W. Yund, "Aided and unaided speech perception by older hearing impaired listeners," *PloS one*, vol. 10, no. 3, p. e0114922, 2015.

[5] B. Kowalewski, J. Zaar, M. Fereczkowski, E. N. MacDonald, O. Strelcyk, T. May, and T. Dau, "Effects of slow-and fast-acting compression on hearing-impaired listeners' consonant–vowel identification in interrupted noise," *Trends in Hearing*, vol. 22, p. 2331216518800870, 2018.

[6] J.-l. Zhang, S.-q. Qi, and S.-n. lü , "A preliminary study of the perceptual configurations of chinese consonants," *Acta Psychologica Sinica*, vol. 1, pp. 76–85, 1982.

[7] J. Zhang, *Fundamentals of Chinese Human-Machine Speech Communication (汉语人机语音通信基础, in Chinese) p(132)*. Shanghai Science and Technology Press, 2010.

[8] S. Zhu, L. L. Wong, F. Chen, and Y. Chen, "Consonant discrimination by mandarin-speaking children with prelingual hearing impairment," *International journal of pediatric otorhinolaryngology*, vol. 79, no. 8, pp. 1354–1361, 2015.

[9] Q. Liu, N. Zhou, R. Berger, D. Huang, and L. Xu, "Mandarin consonant contrast recognition among children with cochlear implants or hearing aids and normal-hearing children," *Otology & Neurotology*, vol. 34, no. 3, pp. 471–476, 2013.

[10] W.-H. Tseng, D.-L. Hsieh, W.-T. Shih, and T.-C. Liu, "Extended bandwidth nonlinear frequency compression in mandarin-speaking hearing-aid users," *Journal of the Formosan Medical Association*, vol. 117, no. 2, pp. 109–116, 2018.

[11] X. Chen, Y. You, J. Yang, J. Qian, Q. Lu, V. Kuehnel, J. Rehmann, B. Liu, and L. Xu, "Effects of nonlinear frequency compression on mandarin speech and sound-quality perception in hearing-aid users," *International Journal of Audiology*, vol. 59, no. 7, pp. 524–533, 2020.

[12] S. Qi, X. Chen, J. Yang, X. Wang, X. Tian, H. Huang, J. Rehmann, V. Kuehnel, J. Guan, and L. Xu, "Effects of adaptive non-linear frequency compression in hearing aids on mandarin speech and sound-quality perception," *Frontiers in Neuroscience*, p. 1038, 2021.

[13] J. Yang, J. Qian, X. Chen, V. Kuehnel, J. Rehmann, A. von Buol, Y. Li, C. Ren, B. Liu, and L. Xu, "Effects of nonlinear frequency compression on the acoustic properties and recognition of speech sounds in mandarin chinese," *The Journal of the Acoustical Society of America*, vol. 143, no. 3, pp. 1578–1590, 2018.

[14] Z. Guo, G. Yu, H. Zhou, X. Wang, Y. Lu, and Q. Meng, "Utilizing true wireless stereo earbuds in automated pure-tone audiometry," *Trends in Hearing*, vol. 25, p. 23312165211057367, 2021.

[15] S. A. Phatak, Y.-s. Yoon, D. M. Gooler, and J. B. Allen, "Consonant recognition loss in hearing impaired listeners," *The Journal of the Acoustical Society of America*, vol. 126, no. 5, pp. 2683–2694, 2009.

[16] B. C. Moore, "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," *Ear and hearing*, vol. 17, no. 2, pp. 133–161, 1996.

[17] B. E. Walden, K. W. Grant, and M. T. Cord, "Effects of amplification and speechreading on consonant recognition by persons with impaired hearing," *Ear and hearing*, vol. 22, no. 4, pp. 333–341, 2001.

[18] S. Moradi, B. Lidestam, M. Hällgren, and J. Rönnberg, "Gated auditory speech perception in elderly hearing aid users and elderly normal-hearing individuals: Effects of hearing impairment and cognitive capacity," *Trends in Hearing*, vol. 18, p. 2331216514545406, 2014.

[19] A. Simpson, A. A. Hersbach, and H. J. McDermott, "Improvements in speech perception with an experimental nonlinear frequency compression hearing device," *International journal of audiology*, vol. 44, no. 5, pp. 281–292, 2005.