Published in final edited form as:

Int J Pediatr Otorhinolaryngol. 2018 October ; 113: 124–130. doi:10.1016/j.ijporl.2018.07.039.

Spoken word recognition in noise in Mandarin-speaking pediatric cochlear implant users

Cuncun Ren^{a,b}, Jing Yang^c, Dingjun Zha^a, Ying Lin^a, Haihong Liu^d, Ying Kong^b, Sha Liu^b, and Li Xu^e

^aDepartment of Otolaryngology, Xijing Hospital, Xi'an, China

^bBeijing Tongren Hospital, Beijing Institute of Otolaryngology, Capital Medical University, Beijing, China

^cCommunication Sciences and Disorders, University of Wisconsin-Milwaukee, Milwaukee, WI, USA

^dBeijing Children's Hospital, Capital Medical University, Beijing, China

^eCommunication Sciences and Disorders, Ohio University, Athens, OH, USA

Abstract

Objective: The purpose of the present study was to compare spoken word recognition performance in the presence of speech spectrum-shaped noise and four-talker babbles in Mandarin-speaking children with cochlear implants (CIs).

Methods: Participants included 33 children with unilateral CIs (with a mean age of 10.4 ± 2.9 years old and a mean length of CI use of 7.5 ± 3.0 years). The Standard Chinese version of Lexical Neighborhood Test was implemented in quiet, speech-spectrum-shaped noise (SSN), and four-talker babble (FTB). The signal-to-noise ratios (SNRs) were set at +5 and +10 dB for both types of maskers. Participants responded by verbally repeating each word they heard and the response was scored as the percentage accuracy of recognition performance. A Generalized Linear Model (GLM) fitting, correlational tests, and a two-way repeated-measures ANOVA were conducted on the percent-correct data.

Results: Word recognition in quiet was on average 74.5% correct but dropped to 57.3% and 48.8% correct for SSN and FTB at 10 dB SNR, respectively, and 44.4% and 32.6% correct for SSN and FTB at 5 dB SNR, respectively. In both quiet and noise conditions, the participants showed lower recognition accuracy for the hard words than for the easy words. Disyllabic words were recognized with higher accuracy rates than were the monosyllabic words. The GLM analysis revealed that all four tested factors (masker type, SNR, lexical neighborhood feature, and lexical type) showed significant impacts on word recognition in children with CIs. Word recognition scores in the two types of maskers were significantly correlated for the disyllabic words at both SNRs and monosyllabic words at 10 dB SNR.

Conclusions: The present study demonstrated that the lexical features such as the lexical neighborhood characteristics and lexical type had significant effects on speech recognition performance in both quiet and noise conditions in pediatric CI users. Children with years of experience of CI use still encountered remarkable difficulties in everyday listening environment although their speech recognition in quiet reached relatively desired level. Fluctuating noise, such as speech babbles, caused greater challenge than steady-state noise for speech recognition in children with CIs.

Keywords

cochlear implant; children; word recognition; noise; Mandarin Chinese

1. Introduction

For patients with severe to profound hearing loss, cochlear implants (CIs) have become the most effective rehabilitation tool. Children with CIs have unprecedented access to important speech information that allows them to develop spoken language skills. As a result, CIs have dramatically improved speech perception ability for hearing-impaired children[1–6]. In recent years, many children with CIs have received education in mainstream schools. However, their auditory, linguistic, and cognitive outcome measures still show great variability in individual performances.

Previous research has shown that CI recipients have reasonably good performance of speech perception in quiet[7]. However, natural acoustic environments often contain multiple sources of noises. Speech recognition tests in noise offer more representative outcomes of perceptual abilities in everyday listening environment for those individuals using auditory devices. Researchers found that hearing-impaired listeners presented considerable deficits in speech recognition in noisy environment[8–12]. For example, Dorman et al.[8] compared the sentence recognition in quiet, noise at 5 and 10 dB signal-to-noise ratios (SNRs) in 15 adults fitted with a CI in one ear and hearing aid in the contralateral ear. These patients were tested using electrical stimulation, acoustic stimulation, and combined stimulation. The results revealed that in all three conditions, the patients showed much lower sentence recognition accuracy in noise than in quiet and the recognition accuracy decreased as the SNR decreased from 10 to 5 dB. In a more recent study, Caldwell and Nittrouer[10] tested speech recognition in quiet and noise at –3, 0, +3 dB SNR in children with normal-hearing (NH), hearing aids (HA), or CIs. They found that all three groups of children showed poorer performance in noise than in quiet. In addition, when the SNR increased from –3 to +3 dB, all three groups of children demonstrated improved recognition accuracy and the magnitude of improvement was the greatest in the children with NH but the smallest in the children with CIs.

Speech spectrum-shaped noise and multi-talker babbles are most frequently used noise types in assessing individual speech recognition ability. When listeners are exposed in noise, two types of masking effects are usually involved. One type of masking effect is energetic masking which is also referred as “physical masking” that occurs when the neural excitation at the auditory periphery evoked by the masker exceeds the excitation produced by the target

speech[13]. This masking effect is determined by the relative energies of the signal and the masker. By contrast, informational masking is one form of “perceptual masking” that can be “broadly defined as a degradation of auditory detection or discrimination of a signal embedded in a context of other similar sounds”[14]. Speech-shaped noise, a continuous steady-state noise, introduces interference of masker at the level of peripheral auditory processing and produces primarily energetic masking. Multi-talker babble, a fluctuant speech noise, introduces informational masking in addition to energetic masking because it is synthesized by mixing connected meaningful speech produced by a number of different speakers. The audible fragments presented in the speech babble masker interfere with a listener’s ability to segregate and/or selectively attend to target speech. The effect of informational masking for speech recognition decreases as the number of talkers increases[15, 16] because the acoustic waveform of the resulting masker babble is less confusable with target speech. In this situation, the masking effect is primarily caused by energetic masking[17]. On the other hand, a masking babble with fewer talkers produces masking release in NH listeners because NH listeners can benefit from the low-intensity parts of the competing noise to recognize speech information. Such ability is referred to as glimpse or listening in dips. However, listeners with hearing loss usually show very little benefit from these dips when tested in the same conditions in comparison to listeners with NH[18].

A few previous studies tested the recognition of pitch and disyllabic words presented with a two-talker babble or a steady-state white noise in adult users of CIs[9,12, 19]. Yet, little data has been reported on the performance of speech recognition in different types of noise in children with CIs. Mao and Xu[20] reported lexical tone recognition in noise in both NH children and prelingually-deafened children with CIs. While children with CIs showed remarkably reduced tone recognition in noise, NH children also performed more poorly than NH adults in adverse listening conditions. This result was consistent with many other reports that children required a more optimal SNR to achieve a reasonably good performance in speech perception than did adults in the presence of various types of background noise[22,23]. The performance gap between children and adults appeared to be larger when the background noise was competing speech babble rather than the steady-state noise[22,24]. Given the different masking effects introduced by speech-spectrum-shaped noise and multi-talker babbles as well as children-adults differences in speech recognition in noise, the present study was designed to examine the performance of speech recognition in different types of noise in children with CI. Specifically, the research aim of the present study was two-folds: first, to determine how the factors of lexical density characteristics, lexical type, noise type, and SNR affect the spoken word recognition performance in children with CIs; and second, to examine whether children with CIs perform differently in steady-state speech-shaped noise and multi-talker speech babbles.

To address these research aims, a Standard-Chinese version of Lexical Neighborhood Test (LNT) was used in the present study. LNT was developed under the theoretical framework of Neighborhood Activation Model (NAM) which states that a stimulus word is recognized in a context of competing words activated in the memory that show similar acoustic-phonetic representations to the target word. The selection process is affected by the word frequency of the stimulus, the density and word frequency of the neighbors. LNT, based on NAM, was

developed with child-friendly words selected from the corpus of Child Language Data Exchange System (CHILDES). This test includes two types of words: lexically easy words and lexically hard words. The former includes those with high word frequency and low neighborhood density and the latter includes those with low word frequency and high neighborhood. Following the same principles of LNT, a Standard-Chinese version LNT was developed. It includes monosyllabic and disyllabic words each further divided into monosyllabic-easy (ME) and monosyllabic-hard (MH), and disyllabic-easy (DE) and disyllabic-hard (DH) words. Details regarding the development of Standard-Chinese LNT and the word density and frequency were provided in Liu et al.[25].

Liu et al.[2] used Standard-Chinese LNT to test prelingually-deafened Mandarin-speaking children with CIs in quiet condition. This study included 230 pediatric Mandarin-speaking CI users with a wide range of chronological age and length of implant use. The results showed that the children with CIs showed higher accuracy for the easy words than for the hard words. Participants' age at implantation and length of device use also affected their recognition performance. Specifically, children with early implantation age and longer duration of device use showed higher recognition accuracy. However, compared with NH controls, the CI group still showed evident deficit in word recognition.

It is well established that preschoolers and school-aged children require a more desirable SNR than adults to achieve similar performance for speech recognition in the presence of steady-state noise and multi-talker babbles[24]. Ren et al.[26] used Standard-Chinese LNT to test word recognition of NH Mandarin-speaking children in quiet and in speech-spectrum-shaped noise at 0 dB SNR. The authors found that the speech-spectrum-shaped noise caused decreased recognition accuracy in comparison to quiet condition. In both quiet and noise conditions, the NH Mandarin-speaking children showed higher word recognition scores for the easy words than for the hard words. In addition, the recognition accuracy in noise condition increased as a function of the participants' chronological age.

Based on the findings on word recognition in children with CIs in quiet and children with NH in noise[2,26], it is reasonable to predict that children with CIs will perform worse in noisy environment than in quiet condition for word recognition and the recognition accuracy will decrease as the SNR decreases. In addition, the lexical characteristics of the stimulus words are expected to affect the performance of children with CIs in a similar way to the children with NH. That is, the easy words should be recognized with higher accuracy than the hard words. As the present study included two types of noise (speech spectrum-shaped and multi-talker babble) that employ different masking effects, it was of particular interest to determine to what extent the two types of noise affect word recognition in children with CIs.

2. Methods and Materials

2.1. Subjects

Participants in the present study consisted of 33 native Mandarin-speaking children with CIs aged between 6 and 16 years old (mean $\pm$ SD, 10.34 ± 2.96 years). All participants met the candidate criteria: (1) had at least 9 months of CI experience (31 of them had more than 4-years of CI experience); (2) used Mandarin as the primary means of oral communication; (3)

had bilateral severe-to-profound sensorineural hearing impairment and received unilateral cochlear implantation; (4) had no other disabilities besides hearing impairment (e.g., cognitive disabilities, intellectual delays, physical disabilities, visual disabilities, etc.); and (5) were enrolled in mainstream primary schools or kindergartens. The participant demographic characteristics are listed in Table 1. With the CI activated, soundfield threshold of the implanted ear at 0.25, 0.5, 1, 2, and 4 kHz were on average 31.5, 30.0, 30.2, 27.5, and 29.3 dB HL, respectively. The use of human subjects was reviewed and approved by the Institutional Review Board of Beijing Tongren Hospital.

2.2 Materials

The Standard Chinese LNT was used to assess the open-set word recognition in the children with CIs. All stimulus words were drawn from a database containing verbal productions of typically developing children ages 3–6 years old[25]. These words represent early-acquired vocabulary of Mandarin-speaking children and are appropriate to assess open-set spoken word recognition skills in this population. There are 5 disyllabic word lists and 5 monosyllabic word lists. Each list contains 24 items.

Two types of maskers were used in the word-recognition tests: speech-spectrum-shaped noise (SSN) and four-talker babble (FTB). The SSN was generated by deriving the long-term spectrum of all speech tokens in Standard Chinese LNT and then using it to filter a white noise so that the filtered noise had the same spectrum of the speech samples. The FTB was generated by mixing the root-mean-square-equalized text-reading speech produced by four Mandarin-speaking adults (two males and two females).

2.3 Procedures

All tests were conducted in a sound booth with a background noise measured at <20 dB (A). The loudspeaker for the sound field presentation for the CI users was calibrated via a sound field speaker system of the GSI audiometer. The participants were seated at the calibration point, 0° azimuth and 1 m from the loudspeaker. A computer was used to control the stimulus presentations via a GSI-61 audiometer. A custom MATLAB program was used to present the stimuli in quiet and noise at various SNRs. Each participant was tested with disyllable and monosyllable words in quiet, SSN, and FTB conditions. For both SSN and FTB conditions, the listeners were tested at the 10 dB SNR condition first, which was followed by the 5 dB SNR condition. The speech tokens were fixed at 60 dB SPL and noise level was adjusted for the desired SNR and mixed with the speech signal digitally. Participants responded by verbally repeating each word they had heard and were encouraged to make their best guess. Since most of the CI children had more than 4 years of CI experiences and their speech intelligibility was fairly high, we found that the word recognition performance was minimally confounded by their speech production ability. In addition, to ensure accurate judgment of their responses, the children's responses were recorded with a digital recorder and were scored by two raters. If there were inconsistencies between the two raters, they referred back to the recorded responses and discussed until agreement was reached. The test lasted for approximately 30 mins for each participant.

2.4 Data analysis

For each participant in each condition, the result was scored as the percentage of the number of correctly identified words out of the total number of tested words. The percentages of word recognition were treated as binomial data, which were converted using logit transformation[27]. A Generalized Linear Model (GLM) was then performed to examine the effects of lexical neighborhood characteristics (easy vs. hard words), lexical type (monosyllabic vs. disyllabic words), SNR (quiet, 10 dB and 5 dB SNR), and noise type (FTB and SSN) on the recognition performance. Then, a two-way repeated measures ANOVA test was used to examine the effect of noise type and SNR on the recognition performance. Both factors (noise type and SNR) were treated as within-subject factors. The percentage scores were collapsed across easy and hard words as well as disyllabic and monosyllabic words. A rationalized arcsine transformation was applied to the percentage data prior to the ANOVA test[28]. Finally, correlation tests were conducted to examine the relationship of word recognition performance between SSN and FTB conditions.

3. Results

Figure 1 shows the percentages of word recognition accuracy for easy words and hard words in disyllables and monosyllables in quiet, SSN and FTB conditions. First, the participants showed lower recognition scores for the hard words than for the easy words in all conditions. Second, the disyllabic words were recognized with higher accuracy rates than were the monosyllabic words. Third, for both types of syllables, the word recognition was adversely affected by the masking noise. The listeners demonstrated lower recognition accuracy in both SSN and FTB at both SNR conditions in comparison to quiet condition. For both types of maskers, the children with CIs showed higher word recognition accuracy at 10 dB SNR than at 5 dB SNR. Finally, the children with CIs showed lower accuracy with the FTB masker than with the SSN masker for both monosyllabic and disyllabic words. A GLM fitting with lexical neighborhood characteristics (easy vs. hard words), lexical type (monosyllabic vs. disyllabic words), and SNR (quiet, 10 dB and 5 dB SNR) as the three main factors showed highly significant effects for all three factors (all $p < 0.001$). A separate GLM fitting with noise type (FTB vs. SSN) as the main factor after removing the word-recognition data in the quiet condition also showed a significant effect ($p = 0.0024$).

One important goal of the present study was to compare the speech recognition performance in different types of noise in children with CIs. To directly compare the participants' performance in SSN vs. FTB at different SNRs, the percentage scores were collapsed across easy and hard words as well as monosyllabic and disyllabic words. Note that the group-mean word recognition score in quiet was 74.5% correct. The scores were 57.3% and 48.8% correct for SSN and FTB, respectively, at 10 dB SNR but dropped to 44.4% and 32.6% correct for SSN and FTB at 5 dB SNR (Fig. 2). The recognition differences between the two types of maskers were 8.5 and 11.9 percentage points for 10 and 5 dB SNRs, respectively. A two-way repeated-measures ANOVA yielded significant main effects of noise type ($F(1,32) = 470.45$, $p < 0.001$) and SNR ($F(1,32) = 128.56$, $p < 0.001$). No interaction effect between noise type and SNR was found.

Figures 1 and 2 demonstrated that the recognition performance of the children with CIs varied in a similar pattern with lexical feature and SNR in SSN and FTB. Correlational analyses were then implemented to examine the relationship between these two types of noise in word recognition of the children with CIs. Figure 3 shows the relationship of word-recognition scores between SSN and FTB masking conditions at the two SNRs for the disyllabic and monosyllabic words, respectively. The recognition scores for easy and hard words in the LNT were collapsed for each participant. The correlational tests yielded significant correlations between SSN and FTB for disyllabic words at both SNRs. However, for monosyllabic words, a significant correlation between the two types of noise was found only at 10 dB SNR but not at 5 dB SNR. We further collapsed the word-recognition scores across the two SNR levels and lexical types for each participant. The relationship between word recognition performance in SSN and FTB masking conditions of the collapsed data is shown in Fig. 4. All participants except one performed more poorly in the FTB conditions than in the SSN conditions. A strong correlation existed in word-recognition performance between the two types of maskers ($r = 0.868$, $p < 0.001$).

Discussion

In cochlear implants, speech signals and ambient noise are all into the speech processor through a single microphone. Limited spectral resolution, lack of sound location information, lack of pitch information, coupled with impaired auditory periphery, all contribute to poor speech recognition in noise in CI users. The present study compared spoken word recognition performance in speech-spectrum-shaped noise and four-talker speech babbles in Mandarin-speaking, prelingually-deafened children with CIs. Consistent with the findings of speech recognition in quiet by children with CIs[2], our participants' word recognition performance in noise was affected by the lexical features of the stimuli. That is, they showed poorer recognition performance for the hard words and monosyllabic words than for the easy words and disyllabic words in both types of noise. These results confirmed the effects of lexical neighborhood characteristics and linguistic context on speech perception in noisy environment in children with CIs. Previous studies[26,29] reported that the lexical neighborhood effect of easy vs. hard words exerted a stronger influence in NH children in noisy environment than in quiet condition. Researchers proposed that this was because listeners had limited access to the acoustic-phonetic features of the input signals in noisy environment. Therefore, they relied more on the lexical and linguistic information to recognize the speech stimuli. In the present study, we observed greater performance differences between the easy and hard words in both SSN and FTB at different SNRs than in quiet for the disyllabic words in the children with CIs (Fig. 1). However, the stronger effect of lexical neighborhood characteristics in noise than in quiet was not consistently shown in the monosyllabic words. This might be attributable to the lack of semantic clues in the monosyllabic words in comparison to the disyllabic words. In disyllabic word recognition, the children with CIs could benefit more from the assistance of semantic context in addition to the lexical neighborhood features in noise than in quiet. By contrast, in monosyllabic word recognition, due to the absence of semantic clues, the lexical neighborhood features might cause greater performance differences between easy and hard words in quiet than in noise.

Compared to previous studies examining speech recognition of children with CIs, the present study is informative in that we evaluated the effects of different types of noise on speech recognition in this population. It is noteworthy that the children with CIs performed considerably worse in both types of noise than in quiet, which was not surprising and was consistent with previous findings in both adults and children with CIs [8, 10]. Compared to steady-state speech-shaped noise, multi-talker babbles present greater spectrotemporal fluctuations in masker energy. NH listeners can take advantage of the spectrotemporal minima in the masker to detect target speech signals[30]. Compared with NH listeners, however, hearing-impaired listeners benefit less from the masking release and the facilitating role of masking release decreased as the hearing loss worsened[31]. In the present study, we observed different patterns of speech recognition performance in speech-shaped noise and four-talker babble at different SNRs in Mandarin-speaking children with CIs. The participants performed much worse in FTB than in SSN. This result indicated that multi-talker babble was a stronger masker than steady-state speech-shaped noise, especially in less optimal SNR for the children with CIs. In contrast to previous findings of the beneficial role of fluctuating speech babble noise in speech recognition in NH adults, the performance of word recognition in the children with CIs exacerbated in speech babbles relative to speech-shaped noise. Leibold & Buss[23] examined consonant identification in speech-shaped noise and two-talker babbles by NH children at different age groups in comparison to NH adults. The results showed that while the NH adults performed equally well in speech-shaped noise and two-talker babble at 0 dB SNR, all children performed much worse in two-talker babbles than in speech-shaped noise. Further, the children at a younger age showed a much greater performance differences between the two types of maskers than did the children at an older age. That is, the adverse effect of two-talker babble was stronger in the younger children than in the older children. These findings, together with the poorer recognition performance in four-talker babbles than in speech-shaped noise in the children with CIs, suggest that children, regardless of their hearing status, tend to be detrimentally affected by the multi-talker babbles than by the speech-shaped noise.

We speculate that informational masking may have played a role in this outcome as children tend to show a larger effect of informational masking than adults [24]. The ability to segregate and selectively attend to the target speech signal versus masker may not be well developed in children. In addition, this result may also be partially attributable to the developing cognitive abilities and auditory attention in children. A number of earlier studies have shown that children experience more difficulties in maintaining auditory attention and segregating irrelevant stimulus from the target stimulus[32,33]. The children with CIs were likely to be more “distracted” by and paid more attention to the speech babble maskers than the SSN, especially when the SNR was less favorable which made the babble speech contents more prominent.

Although the speech-shaped noise and multi-talker babbles introduced different masking effects to the children with CIs and caused different outcomes in recognition performance, the word recognition performance under the two types of noise were significantly correlated for disyllabic words at both 10 and 5 dB SNRs and monosyllabic words at 10 dB SNR (Fig. 3). In monosyllabic words at 5 dB SNR, the recognition performance was not significantly correlated between SSN and FTB although they showed a trend of positive correlation. This

might be due to results approaching a floor effect in recognition performance in FTB at the inferior SNR level. These results indicate that children who perform more poorly in speech-shaped noise are likely to perform more poorly in multi-talker babbles, and vice versa. Recently, Corbin et al.[34] tested the recognition of monosyllabic words in two-talker speech babble and speech-shaped noise by three age groups of children and adults. At the group level, the speech reception thresholds in speech-shaped noise and two-talker babbles were clearly correlated with each other. However, it was surprising that at individual level, they found that the speech reception thresholds in the two types of maskers were not significantly correlated to each other. The present study and Corbin et al.[34] study had clear procedural differences; we measured percent-correct scores at fixed SNRs and Corbin et al. study tested speech reception threshold. Secondly, Corbin et al. study employed two-talker babbles and the present study used four-talker babbles. Presumably, two-talker babbles introduced greater informational masking than four-talker babble because of more salient speech information involved with a smaller number of competing talkers. The main difference was probably the subjects; we tested pediatric CI user here and Corbin et al. study recruited NH children. It remained to be determined whether the pediatric CI users employed different perceptual strategies from the NH children for different types of maskers.

Although informative, the present study only had 33 participants who varied in chronological age, age at implantation, and length of CI use. The recognition performance also demonstrated large variability. For future study, a larger number of children with CIs should be recruited. Speech recognition including both word and sentence recognition should be tested. Note that the present study focused on Mandarin-speaking pediatric CI users. Mandarin is a tonal language in which pitch contours convey lexical meanings[35]. Although prosodic cues are beneficial for speech recognition in noise at the sentence level[36,37], such advantages of tonal languages are probably lacking because pitch information is poorly coded in CIs. From this perspective, the recognition performance of Mandarin-speaking children with CIs in different types of noise shown in the present study can be generalized to pediatric CI users from other language backgrounds. Practically, the outcome of the present study bears important clinical implications. As background noise especially multi-talker babbles at low SNRs exerts greater adverse effects in speech recognition of children with CIs, parents, school teachers, and clinicians should provide more support to create an ideal, quiet environment in order for the hearing-impaired children to develop the speech, language, and hearing abilities.

Conclusions

In summary, the present study revealed significant lexical effects (the lexical neighborhood characteristics and lexical type) on word recognition in noise in pediatric CI users whose native language was Mandarin Chinese. Further, children with a relatively long experience of CI use still encountered remarkable difficulties in everyday listening environments, although their speech recognition in quiet reached a relatively desired level. In comparison to speech-shaped noise, four-talker babbles caused greater challenges for speech recognition in children with CIs. The adverse effects of noise on speech recognition in children with CIs worsened in less optimal SNRs by a similar amount for both types of maskers. The word-

recognition scores in the two types of maskers correlated with each other for the disyllabic words at both 5 and 10 dB SNRs and for the monosyllabic words at only 10 dB SNR. Such a correlation became very strong when the word-recognition scores were collapsed across the word type or SNR levels. The Lexical Neighborhood Test in noise is an effective way to evaluate speech recognition performance in everyday situations in pediatric hearing-impaired listeners.

Acknowledgments

This study was partially supported by grants from the National Institutes of Health National Institute on Deafness and Other Communication Disorders (Grant No. R15-DC014587), National Natural Science Foundation of China (Grant No.81170916), and Xijing Promote Medical Service Special Fund (Grant No. XJZT15D02).

References

- [1]. Geers A, Tobey E, Moog J, Brenner C, et al. Long-term outcomes of cochlear implantation in the preschool years: from elementary grades to high school, *Int. J. Audiol* 47(sup2) (2008) S21–S30. [PubMed: 19012109]
- [2]. Liu H, Liu S, Wang S, Liu C, Kong Y, et al. Effects of lexical characteristics and demographic factors on mandarin Chinese open-set Words recognition in children with cochlear implants, *Ear Hear*. 34(2013) 221–228. [PubMed: 23086423]
- [3]. Liu H, Liu S, Kirk KI, Zhang J, Ge W, et al. Longitudinal performance of spoken word perception in Mandarin pediatric cochlear implant users, *Int. J. Pediatr. Otorhinolaryngol* 79(2015) 1677–1682. [PubMed: 26250440]
- [4]. Niparko JK, Tobey EA, Thal DJ, et al. Spoken language development in children following cochlear implantation, *J Am Med Assoc*, 303(15) (2010) 1498–1506.
- [5]. Kronenberger WG, Pisoni DB, Henning SC, et al. Executive Functioning Skills in Long-Term Users of Cochlear Implants: A Case Control Study. *J Pediatr. Psychol* 38(8) (2013) 902–14. [PubMed: 23699747]
- [6]. Tye-Murray N, Spencer L, Woodworth GG, Acquisition of speech by children who have prolonged cochlear implant experience, *J. Speech Hear Res*. 38(2) (1995) 327–337. [PubMed: 7596098]
- [7]. Wilson B, Dorman M, Cochlear implants: A remarkable past and a brilliant future, *Hearing Res*. 242 (2008) 3–21
- [8]. Dorman MF, Gifford RH, Spahr AJ, McKarns SA, The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies, *Audiol. Neurotol* 13(2) (2008) 105–112.
- [9]. Turner CW, Gantz BJ, Vidal C, Behrens A, Henry BA, Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing, *J. Acoust. Soc. Am* 115(4) (2004) 1729–1735. [PubMed: 15101651]
- [10]. Caldwell A, Nitttrouer S, Speech perception in noise by children with cochlear implants, *J. Speech. Lang. Hear. R* 56(1) (2013) 13–30.
- [11]. Chadha NK, Papsin BC, Jiwani S, Gordon KA, Speech detection in noise and spatial unmasking in children with simultaneous versus sequential bilateral cochlear implants, *Otol. Neurotol*, 32(7) (2011) 1057–1064. [PubMed: 21725259]
- [12]. Fetterman BL, Domico EH, Speech recognition in background noise of cochlear implant patients, *Otolaryngol. Head. Neck* 126(3) (2002) 257–263.
- [13]. Jerger J, Editorial: Informational masking, *J. Am. Acad. Audiol*, 17(6) (2006).
- [14]. Leek MR, Brown ME, Dorman MF, Informational masking and auditory attention, *Percept. Psychophys* 50(3) (1991) 205–214. [PubMed: 1754361]
- [15]. Brungart DS, Informational and energetic masking effects in the perception of two simultaneous talkers, *J. Acoust. Soc. Am* 109(2001) 1101–1109. [PubMed: 11303924]

- [16]. Rosen S, Souza P, Ekelund C, Majeed AA, Listening to speech in a background of other talkers: Effects of talker number and noise vocoding, *J. Acoust. Soc. Am* 133 (4) (2013) 2431–2443 [PubMed: 23556608]
- [17]. Freyman RL, Balakrishnan U, Helfer KS, Effect of number of masking talkers and auditory priming on informational masking in speech recognition, *J. Acoust. Soc. Am* 115(2004) 2246. [PubMed: 15139635]
- [18]. Lorenzi C, Husson M, Ardoint M, Debrulle X, Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception, *Int. J. Audiol* 45(2006) 487–495. [PubMed: 17005491]
- [19]. Gfeller K, Turner C, Oleson J, Zhang X, Gantz B, et al. Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise, *Ear Hear.* 28(3) (2007) 412–423. [PubMed: 17485990]
- [20]. Mao Y, Xu L, Lexical tone recognition in noise in normal-hearing children and prelingually deafened children with cochlear implants, *Int. J. Audiol* 56(2017) 23–30.
- [21]. Wilson RH, Farmer NM, Gandhi A, Shelburne E, Weaver J, Normative data for the Words-in-Noise Test for 6- to 12-year-old children, *J. Speech. Lang. Hear. Res* 53(2010) 1111–1121. [PubMed: 20699343]
- [22]. Bonino AY, Leibold LJ, Buss E, Release from perceptual masking for children and adults, *Ear Hear.* 34(2013) 3–14 [PubMed: 22836239]
- [23]. Leibold LJ, Buss E, Children's identification of consonants in a speech-shaped noise or a two-talker masker, *J. Speech. Lang. Hear. Res* 56(2013) 1144–1155. [PubMed: 23785181]
- [24]. Hall JR, Grose JH, Buss E, Dev MB, Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children, *Ear Hear.* 23(2002) 159–165. [PubMed: 11951851]
- [25]. Liu C, Liu S, Zhang N, Yang Y, Kong Y, et al. Standard-Chinese Lexical Neighborhood Test in normal-hearing young children, *Int. J. Pediatr. Otorhinolaryngol* 75(2011) 774–781. [PubMed: 21458862]
- [26]. Ren C, Liu S, Liu H, Kong Y, Liu X, et al. Lexical and age effects on word recognition in noise in normal-hearing children, *Int. J. Pediatr. Otorhinolaryngol* 79(12) (2015) 2023–2027. [PubMed: 26545791]
- [27]. Warton DI, Hui FK, The arcsine is asinine: The analysis of proportions in ecology, *Ecology.* 92(2011) 3–10. [PubMed: 21560670]
- [28]. Studebaker GA. A rationalized arcsine transform, *J. Speech Hear. Res* 28(1985) 455–462. [PubMed: 4046587]
- [29]. Krull V, Choi S, Kirk KI, Prusick L, French B, Lexical effects on spoken-word recognition in children with normal hearing, *Ear Hear.* 31 (1) (2010) 102–114. [PubMed: 19701087]
- [30]. Cooke M, A glimpsing model of speech perception in noise, *J. Acoust. Soc. Am* 119(3) (2006) 1562–1573. [PubMed: 16583901]
- [31]. Bacon SP, Opie JM, Montoya DY, The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds, *J. Speech. Lang. Hear. Res* 41(3) (1998) 549–563. [PubMed: 9638921]
- [32]. Berman S, Friedman D, The development of selective attention as reflected by event-related potentials, *J. Exp. Child Psychol* 59(1995) 1–31. [PubMed: 7876768]
- [33]. Doyle A, Listening to distraction: A developmental study of selective attention, *J. Exp. Child Psychol* 15(1973) 100–115. [PubMed: 4706960]
- [34]. Corbin NE, Bonino AY, Buss E, Leibold LJ, Development of open-set words recognition in children: speech-shaped noise and two-talker speech maskers, *Ear Hear.* 37(2016) 55–63. [PubMed: 26226605]
- [35]. Xu L, Zhou N, Tonal languages and cochlear implants, In Zeng FG, Popper AN, Fay RR(Eds), *Auditory Prostheses: New Horizons*, Springer Science Business Media, LLC, New York, 2011, pp. 341–364.
- [36]. Chen F, Wong LL, Hu Y, Effects of lexical tone contour on Mandarin sentence intelligibility, *J. Speech. Lang. Hear. Res* 57(1) (2014) 338–345. [PubMed: 24023383]

- [37]. Zhu S, Wong LN, Wang B, Chen F, Assessing the importance of lexical tone contour to sentence perception in mandarin-speaking children with normal hearing, *J. Speech. Lang. Hear. Res* 60(7) (2017) 2116–2123. [PubMed: 28672378]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure. 1.

Box plots showing the recognition accuracy for the easy words and hard words in monosyllables (right panel) and disyllables (left panel) in quiet, SSN, and FTB at different SNRs. The white boxes represent recognition of the easy words and the gray boxes represent the recognition of the hard words. The three horizontal lines of each box represent the 25 percentile, mean, and 75 percentile of the percentage scores. The dot in each box represents the median of the data. The whiskers represent the range and the crosses represent outliers of the data.

Figure 2.

Group mean word-recognition performance in SSN (white bars) and FTB (black bars) at 10 and 5 dB SNRs. The error bar represents 1 standard deviation. The dashed line represents the group mean word-recognition performance in quiet at 74.5% correct.

Figure 3.

Relationship of word-recognition scores between SSN and FTB masking conditions for disyllabic words (top panel) and monosyllabic words (bottom panel). The left and right panels represent 10 and 5 dB SNR, respectively. Each dot represents one participant and the solid line is the linear fit of all data ($N = 33$). The correlation coefficient and the p value are shown in the upper left corner of each panel.

Figure 4.

Relationship of word-recognition scores between SSN and FTB masking condition. Each dot represents one participant and the solid line is the linear fit of all data ($N = 33$). The correlation coefficient and the p value are shown in the upper left corner.

Table 1.

Demographic information of the CI participants

Subject #	Sex	Etiology	Age at Implantation (yrs)	Age at Test (yrs)	Duration of CI use (yrs)	Ear implanted
1	M	unknown	6.14	6.96	0.82	L
2	M	unknown	3.04	8.23	5.18	L
3	F	LVAS	1.55	7.80	6.25	R
4	F	unknown	1.33	8.24	6.91	R
5	M	unknown	1.75	7.88	6.13	L
6	M	unknown	3.09	6.56	3.47	L
7	M	LVAS	1.78	12.54	10.76	L
8	F	unknown	2.91	10.95	8.04	L
9	M	unknown	1.78	14.00	12.22	R
10	M	unknown	4.21	12.13	7.92	L
11	M	unknown	1.83	12.62	10.79	L
12	M	unknown	2.88	9.51	6.64	R
13	M	unknown	1.28	7.52	6.24	L
14	F	unknown	2.60	10.60	8.00	L
15	M	LVAS	3.37	9.18	5.82	L
16	M	unknown	2.49	7.66	5.16	L
17	F	LVAS	1.53	8.78	7.25	R
18	F	unknown	1.82	6.11	4.28	L
19	M	unknown	6.39	12.77	6.38	L
20	F	unknown	4.25	12.22	7.96	L
21	F	unknown	6.13	16.37	10.24	R
22	F	unknown	2.08	11.51	9.43	L
23	M	LVAS	1.01	7.32	6.32	R
24	M	unknown	1.53	14.25	12.73	L
25	F	unknown	1.74	14.72	12.98	L
26	F	ototoxicity	2.68	8.44	5.76	L
27	M	unknown	2.64	7.27	4.63	L
28	M	unknown	1.64	13.75	12.11	L
29	M	ototoxicity	2.55	8.43	5.89	L
30	F	unknown	1.98	9.00	7.02	L
31	M	unknown	2.11	13.13	11.01	L
32	F	LVAS	2.27	12.49	10.22	R
33	M	unknown	3.46	5.84	2.38	R

LVAS – Large vestibular aqueduct syndrome