A Cross-Language Comparison of Sentence Recognition Using American English and Mandarin Chinese HINT and AzBio Sentences

Li Xu,[1,4] Xin Xi,[2,4] Alexa Patton,[1] Xianhui Wang,[1] Beier Qi,[3] and Lauren Johnson[1]

Objectives: The aim of this study was to perform a cross-language comparison of two commonly used sentence-recognition materials (i.e., Hearing in Noise Test [HINT] and AzBio) in American English (AE) and Mandarin Chinese (MC).

Designs: Sixty normal-hearing, native English-speaking and 60 normal-hearing, native Chinese-speaking young adults were recruited to participate in three experiments. In each experiment, the subjects were tested in their native language. In experiments I and II, noise and tone vocoders were used to process the HINT and AzBio sentences, respectively. The number of channels varied from 1 to 9, with an envelope cutoff frequency of 160 Hz. In experiment III, the AE AzBio and the MC HINT sentences were tested in speech-shaped noise at various signal to noise ratios (i.e., −20, −15, −10, −5, and 0 dB). The performance-intensity functions of sentence recognition using the two sets of sentence materials were compared.

Results: Results of experiments I and II using vocoder processing indicated that the AE and MC versions of HINT and AzBio sentences differed in level of difficulty. The AE version yielded higher recognition performance than the MC version for both HINT and AzBio sentences. The type of vocoder processing (i.e., tone and noise vocoders) produced little differences in sentence-recognition performance in both languages. Incidentally, the AE AzBio sentences and the MC HINT sentences had similar recognition performance under vocoder processing. Such similarity was further confirmed under noise conditions in experiment III, where the performance-intensity functions of the two sets of sentences were closely matched.

Conclusions: The HINT and AzBio sentence materials developed in AE and MC differ in level of difficulty. The AE AzBio and the MC HINT sentence materials are similar in level of difficulty. In cross-language comparative research, the MC HINT and the AE AzBio sentences should be chosen for the respective language as the target sentence-recognition test materials.

Key words: American English, Cross-linguistic comparison, Mandarin Chinese, Noise masking, Sentence recognition, Vocoder processing.

(Ear & Hearing 2021;42;405–413)

INTRODUCTION

Sentence recognition in quiet and in noise has become an integral part of the clinical evaluation of outcomes in hearing-impaired listeners who are fitted with various hearing devices (e.g., hearing aids and cochlear implants [CI]). In 1994, an American English version of Hearing in Noise Test (AE HINT) sentences was developed (Nilsson et al. 1994). The AE HINT sentences were derived from the original British version of Barnford–Kowal–Bench sentences (Bench & Barnford 1979). They consisted of 26 phonemically balanced lists of 10 sentences. The sentences used children's vocabulary and were at approximately a first-grade reading level. The sentences were recorded using a male talker. The AE HINT sentences have gained widespread use in the clinics in the United States and were adopted in previous versions of the Minimum Speech Test Battery (MSTB) for postlingually deafened adult CI recipients (Luxford et al. 2001). Since then, they have been used in numerous clinical research studies (e.g., Sargent et al. 2001; Skinner et al. 2002; Xu et al. 2005; Gifford et al. 2008, 2015; Giguère et al. 2008; Massa & Ruckenstein 2014).

The prominence of AE HINT led to the development of versions of HINT in many different languages, including Spanish, Mandarin Chinese (MC), Cantonese, Japanese, Norwegian, Turkish, and Farsi, etc. (Soli et al. 2002; Wong & Soli 2005; Wong et al. 2007; Soli & Wong 2008; Shiroma et al. 2008; Myhrum et al. 2016; Darouie et al. 2020). The MC version of HINT (MC HINT) (Wong et al. 2007) contained 12 test lists and 2 practice lists. Each list consisted of 20 sentences, and each sentence was 10 characters long. Chinese character is the basic unit in Chinese writing system. Each character is a syllable, and one or more (usually two) characters make up a word. The sentences were recorded from one male talker. A comparison of the speech reception thresholds (SRTs) of the AE HINT and MC HINT sentences in normal-hearing listeners revealed that both versions of the sentences produced similar SRTs (<1 dB) in quiet (Wong et al. 2007) but differed by 1.7 dB in the noise-front condition (Soli & Wong 2008). In a subsequent study, when listeners with more diverse Chinese dialects were tested, their group mean SRT was only 1.1 dB lower than that in the English HINT norms (Xu et al. 2015). A few studies have used MC HINT to evaluate CI outcomes in Mandarin-speaking recipients (e.g., Zhang et al. 2010; Su et al. 2016; Meng et al. 2019).

The HINT materials were developed, validated, and normed primarily for assessment of SRTs in noise using adaptive test procedures to avoid floor and ceiling effects (e.g., Soli & Wong 2008). However, when testing listeners with hearing loss fitted with hearing devices, a certain proportion of them could not complete the test even with a relaxed adaptive rule (Chan et al. 2008; Zhang et al. 2010). For example, Zhang et al. (2010) tested 23 Mandarin-speaking postlingually deafened adults with CIs and found that eight of them (≈35%) could not perform the adaptive procedures of MC HINT with the most relaxed adaptive rule. In clinical evaluation using the AE HINT sentences in quiet at fixed presentation levels, on the contrary,

[1]Division of Communication Sciences and Disorders, Ohio University, Athens, Ohio, USA; [2]Department of Otolaryngology–Head & Neck Surgery, Chinese PLA General Hospital, Beijing, China; and [3]Beijing Tongren Hospital, Beijing Institute of Otolaryngology, Capital Medical University, Beijing, China; [4]These authors are co-first authors who have contributed equally to this work.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

0196/0202/2021/422-405/0 • Ear & Hearing • Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved • Printed in the U.S.A.

<zdoi; 10.1097/AUD.0000000000000938>

clinicians reported that hearing aid and CI users achieved high performances on average. The ceiling effect observed with the AE HINT sentences indicated that AE HINT sentences are "too easy" and do not provide enough difficulty for the assessment of speech intelligibility, especially in quiet (Luxford et al. 2001; Firszt et al. 2009; Gifford et al. 2008). This issue has led to the development of a new material for sentence recognition test: the AE AzBio sentences. The AE AzBio sentences were created with less contextual information and recorded with multiple talkers to overcome the ceiling effects in evaluating sentence recognition in patients with hearing loss and to represent a more realistic listening situation (Spahr & Dorman 2004; Spahr et al. 2007). Note that the AE AzBio materials were designed to measure intelligibility in quiet and not for adaptive SRT measurements in noise as the HINT materials. To accommodate for more testing conditions, the number of lists was increased to 33, with each list consisting of 20 sentences. In addition, the sentence recordings used four talkers (two females and two males), and each spoke at a normal conversational pace and volume and avoided using overly enunciated speech or clear speech, as was used in the HINT sentence materials (Spahr et al. 2012). Numerous studies have also used the AE AzBio sentences for sentence recognition tasks on normal-hearing population and on CI users (e.g., Dorman et al. 2008; Gifford et al. 2008, 2015; Dorman & Gifford 2010, 2017; Zhang et al. 2013). In 2011, replacing the HINT sentences, the AzBio sentences were adopted as the sentence recognition testing materials in the MSTB for adult CI users in the United States (MSTB 2011; Spahr et al. 2012).

To follow suit, an MC version of AzBio (MC AzBio) sentences has recently been developed (Xi et al. 2015). With reference to the structure of English AzBio lists, the initial MC AzBio sentence corpus included 2000 items adapted from current Chinese television programs and social media on present-day adult topics, excluding proper nouns and idioms. The sentences were recorded from four adult native MC–speaking talkers, including two males and two females. Only the 1020 sentences judged with the highest naturalness were processed through a 5-channel vocoder CI simulation. Sentence recognition of the vocoder-processed sentences in 17 normal-hearing adults yield a mean performance score of 76.7% correct (SD, 15.9%). Finally, 640 sentences with performance within ±1 SD of the mean were included in the MC AzBio sentence material that contained 26 test lists and 6 practice lists, with 20 sentences in each list. Although the principles in developing the MC AzBio sentences were the same as those in developing the AE AzBio sentences, the equivalency between the AE and MC versions of AzBio sentences has not been established.

The number of CI users has increased dramatically in China in recent years (Han et al. 2009; Liang & Mason 2013; Li et al. 2017). The total of number of CI recipients in China is estimated to be more than 70,000 as of 2019 (Liu & Yang 2019). Besides the "Big Three" manufacturers (i.e., Cochlear Corporation of Australia, Med-El of Austria, and Advanced Bionics by Sonova Group of Switzerland), the China domestically-manufactured CI devices (Nurotron) have more than 10,000 recipients since the launch in 2011 (Zeng et al 2015; Rebsher et al. 2018). Although a majority of the CI recipients in China are prelingually deafened children, we have seen a dramatic increase in the number of adult recipients in recent years due to enhanced awareness and improved affordability (Yang et al. 2010). Also, since the first cochlear implantation in China 25 years ago, pediatric recipients have grown up into adulthood. There is an urgent need to carry out more clinical research on speech recognition performance of the Mandarin-speaking CI users. MC is a tonal language, and remarkable deficits of tone perception in Mandarin-speaking CI recipients at the monosyllabic level have been documented in numerous studies (e.g., Han et al. 2009; Wang et al. 2011; Xu & Zhou 2011; Zhou et al. 2013, Chen & Wong 2017; Mao & Xu 2017; Liu et al. 2017). How do such deficits in tone perception at the monosyllabic level translate into sentence recognition performance in Mandarin-speaking CI users? How does the sentence recognition performance of Mandarin-speaking CI users compare to that of English-speaking CI users? Given that copious clinical and research data on the sentence-recognition performance in CI users have been accumulated in English-speaking countries, the question at issue is what test materials in MC are appropriate for the cross-linguistic comparison of the clinical data? The answer to this question will also help to provide guidance for selecting appropriate sentence-recognition test materials if a Chinese version of the MSTB for Mandarin-speaking CI recipients is to be established in the future.

Present Study

The purpose of the present study was to perform a cross-language comparison with two existing sets of sentence recognition materials: HINT and AzBio sentences. One way to compare similarities of sentences was to use normal-hearing listeners performing sentence-recognition tasks with vocoder-processed speech materials (e.g., Spahr et al. 2012). Sentence-recognition scores in percent correct were then compared among the sentence lists. The other way to compare similarities of sentences was to test sentence recognition in masking noise (e.g., Soli & Wong 2008; Schafer et al. 2012). Typically, SRTs were compared between test materials or among sentence lists. In the present study, three experiments were designed to compare sentence recognition performance between the AE HINT and the MC HINT sentence materials (experiment I) and that between the AE AzBio and MC AzBio sentence materials (experiment II). In both experiments I and II, vocoder processing was used to degrade the speech signals. Since current speech-processing strategies used in contemporary multichannel CI systems are essentially based on the same principles of channel vocoders (Wilson et al. 1991; Loizou 2006, Wilson 2019), vocoder as an acoustical simulation of CI processing has widely been used to study the effects of CI processing on speech and music perception (e.g., Shannon et al. 1995; Friesen et al. 2001; Xu et al. 2002, 2005; Fu et al. 2004, 2017; Dorman et al. 2005; Qin & Oxenham 2006; Xu & Pfingst 2008; Meng et al. 2019; Everhardt et al. 2020). Therefore, results from these two experiments are particularly relevant to CI research.

In addition, two types of channel vocoders were employed in both experiments I and II: noise-excited vocoder and tone-excited vocoder. The two types of vocoder differ in the carrier of the envelopes. One is a band-passed noise and the other is a sinusoid with a frequency at the center of the band-pass filter. Dorman et al. (1997) found minor differences between tone- and noise-excited vocoders in English sentence perception. Whitmal et al. (2007) reported clearly superior performance with the tone vocoder over noise vocoder in sentence recognition in quiet and noise. They speculated that intrinsic temporal fluctuations of the noise carriers might have interfered with the temporal fluctuations of speech envelope cues. Similarly, Fu et al. (2017) showed that English vowel recognition improved as the bandwidth of the carrier was reduced. Oliver and Gonzalez (2005) also showed a significantly better performance in gender recognition with tone vocoder than with noise vocoder. The authors speculated that the periodic F_0 information in the envelope was presented as sidebands in the tone vocoder but not in the noise vocoder. Chen and Lau (2014) demonstrated that Mandarin sentence recognition performance with sinusoidal carriers was better than noise carriers. However, no differences were found in Mandarin tone recognition between the noise and tone vocoders. The source of superiority of tone vocoder in sentence recognition could not be attributed to potentially better pitch representation in the tone vocoded signals. Therefore, in the present study, we included both types of vocoders for sentence materials in both languages and attempted to determine the effects of vocoder types on sentence recognition in two different languages.

Since experiments I and II produced interesting results that the AE AzBio sentences and the MC HINT sentences yielded similar recognition performance under vocoder processing, a third experiment (experiment III) was designed to test whether these two sets of sentence materials would produce similar recognition performance in noise conditions.

EXPERIMENT I: AMERICAN ENGLISH VERSUS MANDARIN CHINESE VERSIONS OF HINT SENTENCES

Methods

Participants • A group of 20 native English-speaking, normal-hearing adults (4 males and 16 females; age, 23.95 ± 0.64 years) and 20 native MC–speaking, normal-hearing adults (10 males and 10 females; age, 24.84 ± 1.25 years) were recruited to participate in experiment I. To avoid any potential effects of dialects on sentence-recognition performance, all native English-speaking listeners were recruited in Ohio and all native Mandarin-speaking listeners were recruited in Beijing area. A hearing screening was conducted to ensure that all participants had hearing thresholds ≤20 dB HL between 250 and 8000 Hz. Participants had no experience with vocoder-processed speech prior to this study. The use of human subjects was reviewed and approved by the Institutional Review Board of Ohio University.

Stimuli • The AE HINT sentence lists 1 to 20 (10 sentences per list) were used for the test conditions and lists 21 to 26 were used for the practice conditions. To accommodate the required 100 sentences for practice (see Procedure section below), lists 21 to 24 were used repeatedly. The MC HINT sentence lists 1 to 10 (20 sentences per list) were used for the test conditions, whereas lists 11, 12, and Practice Lists 1 and 2 were used for the practice conditions. To accommodate the required 100 sentences for practice (see Procedure section below), Practice List 2 was used repeatedly.

Signal Processing • The vocoder processing similar to that in the study by Shannon et al. (1995) and Xu et al. (2002, 2005) was used to process all HINT sentences. Two types of vocoders (noise and tone) were used. For noise vocoder, speech signal was divided into a number of frequency bands (i.e., 1, 3, 5, 7, or 9) and the envelope of each band was extracted to using half-wave rectification and low-pass filtering at 160 Hz. A 160-Hz envelope cutoff was chosen to ensure best performance possible with the vocoded speech materials because many studies have documented that a low envelope cutoff is sufficient for speech recognition (e.g., Shannon et al. 1995; Xu et al. 2002, 2005; Fu et al. 2004; Kim et al. 2015). The overall bandwidth was from 150 to 5500 Hz, and the bandwidth of each analysis filter was based on estimated equal distance along the basilar membrane (Greenwood 1990). The temporal envelope of each band was then used to modulate a white noise that was band-passed through the same analysis filter. Finally, the modulated noise bands were summed, and the resultant signals were stored on the computer hard disk for presentation. For the tone vocoder, the signal processing was identical to that of the noise vocoder, except that instead of band-passed noise, sinusoids of frequencies equal to the center frequencies of the analysis filters were used as carriers.

Procedure • The vocoder-processed sentences were presented binaurally to participants through Sennheiser HD280 Professional headphones in a double-walled sound booth with noise levels ≤30 dB (A). A custom MATLAB program was created for stimulus presentation. Presentation level of the stimuli was set by participants at their most comfortable listening level. Prior to the formal testing session, a practice session was conducted to help familiarize the participants to the vocoded speech prior to testing. Each participant completed 100 practice sentences (i.e., 10 Sentences × 2 Types of Vocoder × 5 Channel Conditions). The practice condition started with the highest number of channels and progressed downward for each participant. During the practice, feedback was provided. The practice session typically lasted for 20 to 30 minutes. Each sentence could be repeated up to three times.

The test session consisted of 10 conditions (i.e., 2 Types of Vocoder × 5 Channel Conditions). Each condition was tested with 20 sentences. Therefore, a total of 200 sentences (i.e., 20 Sentences × 2 Types of Vocoder × 5 Channel Conditions) were used in sentence recognition. The type of vocoder and the number of channels used for the vocoder processing were counter-balanced among participants (see Supplemental Digital Content 1, http://links.lww.com/EANDH/A695). For the MC HINT, the condition number and sentence list number were consistent. For example, in condition 1, the MC HINT list 1 (20 sentences) was used. For the AE HINT, 2 lists (i.e., list n and list $n + 10$) were used for any one channel condition. For example, in condition 1, list 1 and list 11 of the AE HINT (20 sentences total) were used. During the test, participants were allowed to repeat each sentence up to three times. The rationale behind allowing three repetitions was to ensure the best possible performance for our listeners to recognize the degraded signals within an acceptable time limit. No feedback was provided in the test session. The test session took each participant approximately 45 to 60 minutes to complete.

Results

Sentence recognition performance was scored based on correctly recognized key words in AE HINT sentences or correctly recognized Chinese characters in MC HINT sentences. Figure 1 shows the group mean sentence recognition performance for the AE and MC HINT sentences. For both noise and tone vocoders, the performance in the AE HINT condition was overall higher than that in the MC HINT condition. A generalized linear mixed model (GLMM) (Warton & Hui 2011) was used to examine the effects of the version of HINT sentences (i.e., AE and MC languages), the number of channels, and the type of vocoder (tone and noise vocoders) on sentence-recognition performance. The analyses revealed a significant effect of language (β = 0.86, t = 23.07, $p < 0.001$), number of channels (β =1.29, t =89.43, $p < 0.001$), and type of vocoder (β = 0.26, t = 7.15, $p < 0.001$). Tone vocoder yielded slightly better mean sentence recognition scores than noise vocoder did.

EXPERIMENT II: AMERICAN ENGLISH VERSUS MANDARIN CHINESE VERSIONS OF AZBIO SENTENCES

Methods

Participants • A new group of 20 native English-speaking, normal-hearing adults (7 males and 13 females) of age 22.39 ± 0.78 years and 20 native MC-speaking, normal-hearing listeners (10 males and 10 females) of age 24.51 ± 2.65 years were recruited to participate in experiment II. Other recruitment criteria were the same as in experiment I.

Stimuli • For English, the 15 AE AzBio sentence lists that produced the most consistent recognition performance tested on a group of CI users in a validation study by Spahr et al. (2012) were selected for perception tasks. The first 10 sentence lists (i.e., lists 2, 3, 4, 5, 10, 11, 12, 14, 21, and 26) with the least variability derived from Spahr et al. (2012) were used for the test conditions, while the following five sentence lists (i.e., lists 16, 17, 18, 22, and 24) with the least variability were used in the practice. For MC, 15 sentence lists from the MC AzBio sentence lists were chosen based on the same principles. The 10 lists for test conditions were lists D, E, H, M, N, O, P, Q, S, and T and the 5 lists for practice were lists A, B, C, F, and I.

Signal Processing • The same signal processing as in experiment I was applied in experiment II.

Procedure • All testing conditions were the same as in experiment I. The type of vocoder and the number of channels used for the simulation were counter-balanced among participants. The same assignment table (see Supplemental Digital Content 1, http://links.lww.com/EANDH/A695) was used to assign a participant to a particular test order and sentence list. For both AE and MC, one list of 20 sentences was used for one test condition.

Results

Figure 2 shows the group mean sentence-recognition performance for the AE and MC AzBio conditions. For both noise and tone vocoders, the performance in the AE AzBio condition was overall higher than that in the MC AzBio condition. A GLMM analysis was conducted to examine the effects of the version of the AzBio sentences (i.e., AE and MC), the number of channels, and the type of vocoder on sentence-recognition performance. Results showed that there was a significant effect of language (β = 0.69, t = 22.91, $p < 0.001$), number of channels (β =0.82, t =103.45, $p < 0.001$), and type of vocoder (β = 0.11, t = 3.69, $p < 0.001$) on sentence recognition performance. Tone vocoder yielded a slightly better mean perception score than noise vocoder did.

Results from both experiments I and II were compared within each experiment and across both experiments. First, we compared sentence-recognition performance between different sentence materials in either noise or tone vocoders. Figure 3 (left and middle) shows the group mean differences in performance across all 3-, 5-, and 7-channel conditions between AE HINT and AE AzBio, MC HINT and MC AzBio, AE AzBio and MC HINT, AE HINT and MC HINT, AE AzBio and MC AzBio, and AE HINT and MC AzBio in noise and tone vocoder conditions. Data from 1- and 9-channel conditions were excluded in these comparisons because of the floor and ceiling effects. Interestingly, the smallest absolute mean difference in performance was found between the AE AzBio and the MC HINT sentence materials, regardless of the type of vocoders. Independent-sample t tests were conducted on the rational arcsine units of the recognition scores to compare the performance between the various sentence materials. The recognition difference between the AE AzBio and MC HINT sentences was not statistically significant in the tone-vocoder condition (t (38) = –1.847, p = 0.073). All other recognition differences between sentence materials were statistically significant (all $p < 0.05$). Note that the differences between AE HINT and MC HINT were driven primarily by the performance difference at the

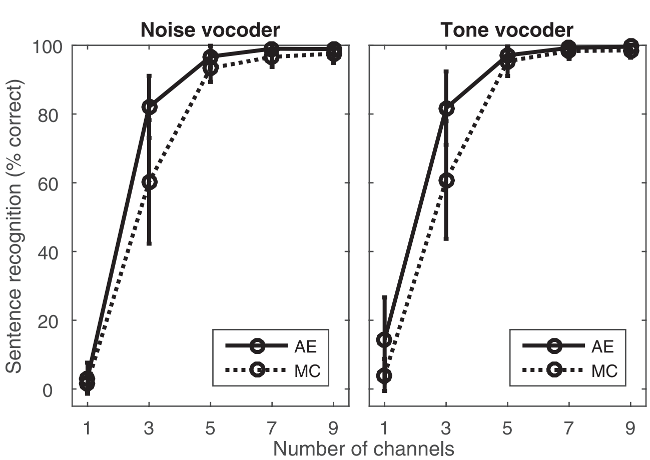

Fig. 1. Group mean Hearing in Noise Test (HINT) sentence recognition scores in noise vocoder and tone vocoder conditions (left and right). In each panel, the solid line represents performance of the American English (AE) HINT condition and the dotted line represents that of the Mandarin Chinese (MC) HINT condition. The error bar indicates ±1 SD.

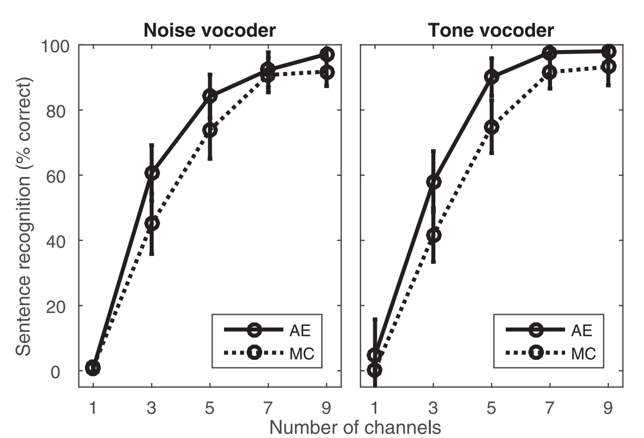

Fig. 2. Group mean AzBio sentence recognition scores in noise vocoder and tone vocoder conditions (left and right). In each panel, the solid line represents performance of the American English (AE) AzBio condition and the dotted line represents that of the Mandarin Chinese (MC) AzBio condition. The error bar indicates ±1 SD.

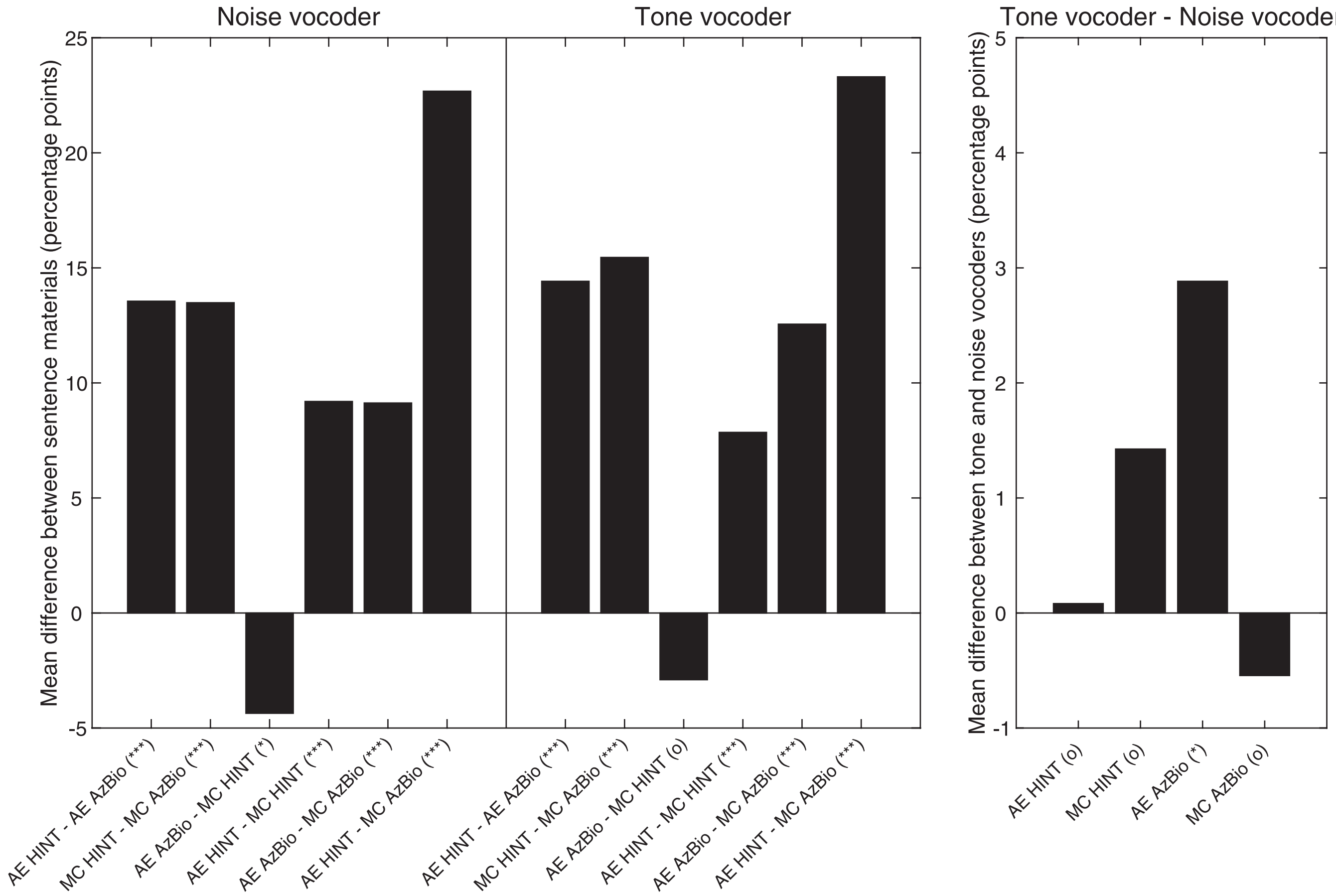

Fig. 3. Left and Middle, Mean differences in sentence-recognition performance between American English (AE) Hearing in Noise Test (HINT) and AE AzBio, Mandarin Chinese (MC) HINT and MC AzBio, AE AzBio and MC HINT, AE HINT and MC HINT, AE AzBio and MC AzBio, and AE HINT and MC AzBio conditions for the noise and the tone vocoder. Right, Mean differences in sentence-recognition performance between tone and noise vocoder in the four sets of sentence materials. The symbols in the parenthesis indicate statistical significance: o, not significant $p > 0.05$, *$p < 0.05$, and ***$p < 0.001$.

3-channel condition because the performance reached plateau at five channels. The AzBio sentences were more difficult than the HINT sentences by design. The mean differences between HINT and AzBio sentences in both languages were approximately 14 percentage points. The differences between HINT and AzBio or between AE and MC AzBio were derived mainly from the three- and five-channel conditions because the performance reached plateau at seven channels. This is consistent with the notion that more channels are required to reach performance plateau for more difficult speech materials (Shannon et al. 2004; Xu & Zheng 2007; Xu & Pfingst 2008).

Next, we compared the sentence-recognition performance in noise versus tone vocoders. The group mean differences in sentence recognition between tone and noise vocoders for the four sets of sentence materials were calculated across the 3-, 5-, and 7-channel conditions. Data from 1- and 9-channel conditions were excluded in these comparisons because of floor and ceiling effects. Results showed that the absolute differences between tone and noise vocoders were within three percentage points for all four sets of sentence materials (0.08 percentage points for the AE HINT, 1.43 percentage points for the MC HINT, 2.88 percentage points for the AE AzBio, and -0.54 percentage points for the MC AzBio sentences) (Fig. 3, right). Independent *t* tests conducted on the rational arcsine unit values of the recognition scores indicated that only the difference in the AE AzBio sentences between tone and noise vocoders was statistically significant ($t(38) = 2.08$, $p = 0.044$).

Since the absolute differences in recognition scores between tone and noise vocoders were on average only 1.23 percentage points across all speech materials, we combined the data of both types of vocoders and plotted the group mean sentence recognition performance in Figure 4. The AE HINT sentences achieved the highest scores and the MC AzBio sentences achieved the lowest scores. The recognition scores for the MC HINT sentences and the AE AzBio sentences were in between and overlapped with each other (Fig. 4). The differences of recognition between the MC HINT sentences and the AE AzBio sentences at 1-, 3-, 5-, 7-, and 9-channel conditions were 0.2, 1.2, 7.2, 2.5, and 0.5 percentage points, respectively.

EXPERIMENT III: AE AZBIO AND MC HINT SENTENCE RECOGNITION IN NOISE

Methods

Participants • A new group of 20 native English-speaking, normal-hearing listeners (2 males and 18 females; age, 23.89 ± 2.38 years) and 20 native MC-speaking, normal-hearing listeners (10 males and 10 females; age, 21.75 ± 1.83 years) were recruited to participate in experiment III. The recruitment criteria were the same as in experiments I and II.

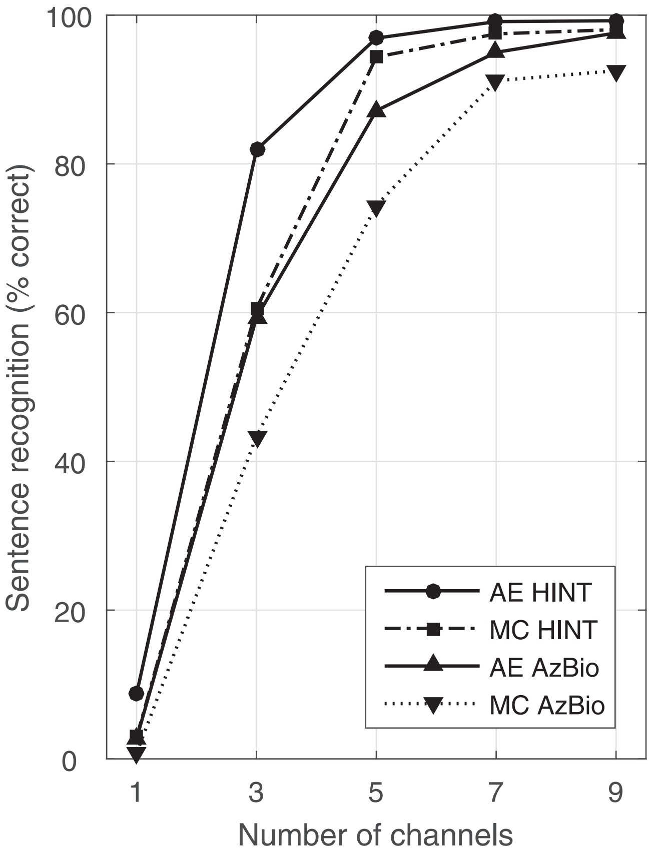

Fig. 4. Group mean sentence recognition scores for the four sets of sentence materials. For each curve, data from noise-vocoder and tone-vocoder conditions were combined. AE, American English; HINT, Hearing in Noise Test; MC, Mandarin Chinese.

Stimuli • Five lists of the AE AzBio sentences (randomly selected from lists 2, 3, 4, 5, 10, 11, 12, 14, 21, and 26) and five lists of the MC HINT sentences (randomly selected from lists D, E, H, M, N, O, P, Q, S, and T) were used in experiment III. The masking noise was the speech-spectrum–shaped noise (SSN) that accompanied the respective sentence materials (Fig. 5). The level of the sentences was fixed and the level of the SSN was adjusted to reach the desired signal to noise ratios (SNRs; i.e., −20, −15, −10, −5, and 0 dB). Before the beginning of the sentence, a 1.4-s long SSN was added. In the middle of

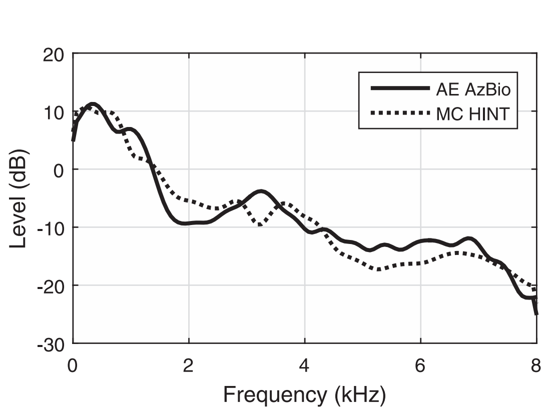

Fig. 5. Spectrum of the speech-shaped noise (SSN) for the American English (AE) AzBio (solid line) and the Mandarin Chinese (MC) Hearing in Noise Test (HINT) sentences (dotted line).

the 1.4-s SSN, a 400-ms long tone burst of 1 kHz was embedded to alert the listeners of the beginning of the sentence. A 0.6-s long SSN was added after the end of the sentence. For the brief practice session, 15 sentences (three sentences at each of the 5 SNRs) were randomly selected from lists 16, 17, 18, 22, and 24 of AE AzBio and lists A, B, C, F and I from MC HINT.

Procedure • The test procedure for experiment III was similar to that of experiments I and II. The test orders of different SNRs were randomized. All participants completed a brief practice session that included 15 sentences (i.e., 3 Sentences × 5 SNRs) to familiarize themselves with the task. The practice condition started with the highest SNR and progressed downward for each participant. During the practice, feedback was provided. The practice session typically lasted <5 minutes. Each sentence could be repeated up to three times.

Results

Figure 6 (left) shows the mean sentence-recognition performance in SSN with the AE AzBio and the MC HINT sentences. In both conditions, the performance improved as the SNR increases at a similar rate. A GLMM analysis revealed significant main effects of SNR ($\beta = 0.49$, $t = 84.03$, $p < 0.001$) and language ($\beta = 0.13$, $t = 3.08$, $p = 0.002$). Although the effect of language was statistically significant, the absolute differences in sentence-recognition performance at SNRs of −20, −15, −10, −5, and 0 dB were only 0.3, 1.1, 0.1, 5.1, and 1.1 percentage points, respectively. The right panel of Figure 6 shows the logistic fitting curves (i.e., the performance-intensity [PI] functions) for the data. The SRTs calculated based on the PI functions were −7.6 and −7.9 dB SNR for the AE AzBio and the MC HINT, respectively. The slopes of the PI functions at 20% to 80% correct recognition were 10.0 and 11.3 percentage points/dB SNR for the AE AzBio and the MC HINT, respectively.

DISCUSSION

In experiments I and II, we compared sentence recognition of HINT and AzBio sentence materials under vocoder processing. Results showed that the HINT sentences yielded higher recognition performance in the normal-hearing listeners under 3-, 5-, and 7-channel vocoder processing than the AzBio sentences in both languages (Fig. 3). These results were not surprising because the AzBio sentences were designed to have a higher level of difficulty than the HINT sentences. This elevated level of difficulty in AE AzBio has been confirmed in adult CI users. Spahr and Dorman (2004) tested a group of 30 adult CI users who were in the upper half of the population of CI recipients in terms of speech perception performance. The results showed that the average recognition scores for the AE HINT and AzBio sentences in quiet were approximately 95% and 75% correct, respectively.

Our results also showed that the AE versions of both sets of sentences produced higher recognition performance in the normal-hearing listeners when compared with the MC versions (Figs. 1 and 2). The AE HINT sentences used children's vocabulary and were at approximately a first-grade reading level. The number of syllables of the AE HINT sentences is usually six to seven, whereas the MC HINT sentences are all 10-character long. For example, a typical AE HINT sentence, "The cat lay on the bed," has six syllables with both articles not counted in

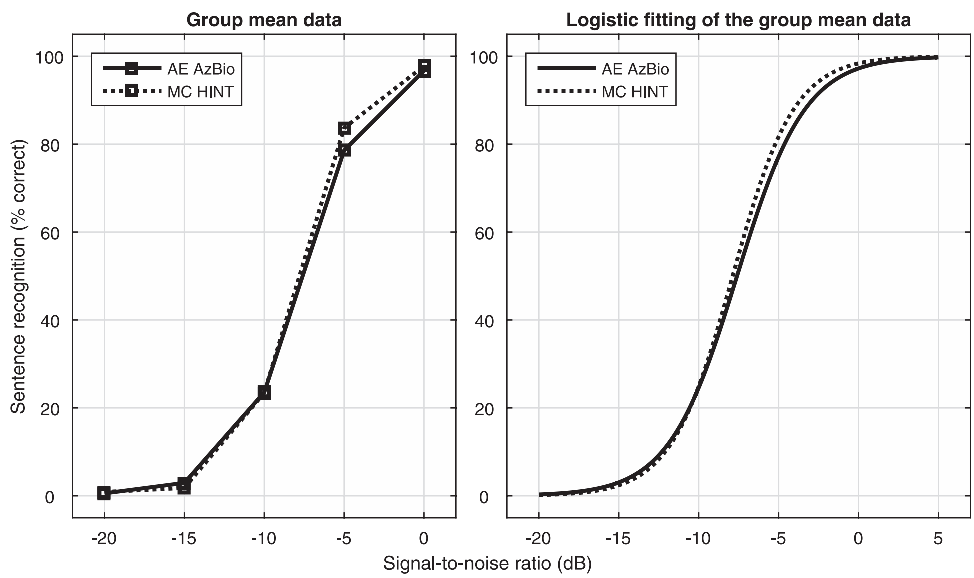

Fig. 6. Group mean sentence-recognition performance in the American English (AE) AzBio (solid line) and the Mandarin Chinese (MC) Hearing in Noise Test (HINT; dotted line) sentences in speech-shaped noise (SSN) at various signal to noise ratios (SNRs; left) and the logistic fitting curves of the corresponding group mean data (right).

scoring. In an MC HINT sentence, “楼下的小猫整晚都在叫” (The little cat downstairs was meowing all night), it seems that the words “little” and “downstairs” are extra in the MC HINT sentence compared to the simpler AE HINT sentence structure. In fact, adjectives and adverbial clauses were added to short sentences to meet the 10-character length requirement for the MC HINT sentences (Wong et al. 2007). Although a detailed linguistic analysis of the HINT sentences in the two languages is beyond the scope of the present study, the MC HINT sentences appear to be more complex at the semantic level and vocabulary level than the AE counterparts. This might have induced an increase in the level of difficulty in the MC HINT sentences. The levels of difficulty in AE AzBio and MC AzBio sentences are not apparently different. They both contain diverse types of sentences with varying demands on vocabulary. The AE AzBio sentences appear to be in a simple subject–verb–object (SVO) structure. For example, “She watched the sneak preview of the new film.” In contrast, the MC AzBio sentences frequently contain noun clauses. For example, “He did not detect that the boss was angry” (他没有察觉到老板生气了). However, this casual observation, without a detailed linguistic analysis of all sentences, should not be taken to fully explain why the AE AzBio sentences produced higher recognition scores than the MC AzBio sentences.

Coincidentally, the AE AzBio sentences and MC HINT sentences yielded similar recognition performance under vocoder processing (Figs. 3 and 4). This result prompted us to conduct experiment III, in which these two set of sentences were used in recognition tasks under noise conditions. Among the SNRs tested in experiment III (i.e., −20, −15, −10, −5, and 0 dB), visible differences were found only at −5 dB SNR (Fig. 6, left). The PI functions of the two sets of sentences were closely matched with each other (Fig. 6, right). The SRTs calculated based on the PI functions differed by only 0.3 dB between the AE AzBio and the MC HINT sentences. Thus, experiment III confirms that the level of difficulties in the AE AzBio and the MC HINT are similar in noise conditions. The SRT in noise for the MC HINT was −7.9 dB SNR in the present study. The SRT value reported by Soli and Wong (2008) in the noise-front condition was −4.3 dB SNR. Direct comparisons of the SRTs should be avoided because the scoring methods and procedural details of the two studies were quite different. In the present study, sentence-recognition performance was based on the percent of words/characters recognized in the sentences, whereas in Soli and Wong (2008) study, performance was calculated based on the percent of sentences correctly recognized. The latter employed a more stringent criterion for scoring and thus yielded a higher SRT.

A second purpose of the present study was to examine the effects of the type of vocoders (i.e., tone and noise vocoders) on sentence recognition in English and in MC. Our results showed that for either language, the differences in sentence recognition with tone vocoder and noise vocoder were very small (Fig. 4, right). These were in agreement with the findings by Dorman et al. (1997) who found minor differences between tone and noise vocoder in speech recognition. Souza and Rosen (2009) found that tone vocoder yielded better speech recognition than noise vocoder when the envelope cutoff frequency was 300 Hz but was less intelligible than noise vocoder for a 30-Hz low-envelope cutoff frequency. Xu (2016) found that sentence recognition using tone vocoder with a 40-Hz envelope cutoff frequency was poorer than using a noise vocoder with a 160-Hz envelope cutoff frequency. Whitmal et al. (2007) used a 300-Hz low-pass cutoff for the envelope extractor in the vocoder processing and found that tone vocoder produced better speech recognition than noise vocoder in noise and in quiet conditions. We used a 160-Hz low-pass cutoff for the envelope extractor in the present study. This intermediate envelope cutoff might be the reason for a similar recognition performance in tone and noise vocoders. In recent studies, Chen and Lau (2014) and Chen et al. (2017) showed that a tone vocoder yielded higher intelligibility scores for Mandarin sentences but not for a lexical tone identification task. Note that their speech signals were corrupted by noise prior to vocoder processing. Thus, when interpreting recognition results using vocoded speech signal, the vocoder parameters such as the envelope cutoff, the type of speech materials, and the noise condition should all be taken into account (Shannon et al. 1995; Xu et al. 2005; Xu & Zheng 2007; Souza & Rosen 2009; Kim et al. 2015).

CONCLUSIONS

The present study tested recognition of AE and MC sentences using vocoder processing in 1- to 9-channel conditions. Speech materials included AE and MC versions of HINT and AzBio sentences. With a 160-Hz envelope low-pass cutoff, no differences in recognition performance were seen between the tone and noise vocoders for either language or for either set of sentence materials. The AzBio sentences were more difficult than the HINT sentences under vocoder processing for both languages. With a large number of sentence lists, the MC AzBio sentence material should be a valuable resource in both clinical and basic research when a great number of test conditions are involved. The AE version of HINT and AzBio sentences produced higher intelligibility scores than the corresponding MC sentences. Coincidently, the AE AzBio sentences and the MC HINT sentences yielded similar recognition performance under vocoder processing. Further testing of sentence recognition in noise conditions confirmed that the two sets of sentences (i.e., AE AzBio and MC HINT) generated similar PI functions. Since the AE AzBio sentences were recommended in the MSTB for CI users in the United States, to facilitate cross-linguistic comparisons of data from MC-speaking CI users, it is recommended that the MC HINT sentences should be used in that population. Clearly, whether this recommendation will stand the clinical validation with CI recipients awaits future studies.

ACKNOWLEDGMENTS

We express our appreciation to Jiaxing Liu and Shizhe Xu for assisting in data collection and to Lexi Neltner and Devon Zurovchak for their editorial assistance in the preparation of the article.

L. X. and X. X. designed the experiments. L. X., X. X., A. P., B. Q., and L. J. performed the experiments. L. X. and X. W. analyzed the data. L. X. and X. X. wrote the main article. All authors discussed the results and implications and commented on the article at all stages.

The authors have no conflicts of interest to disclose.

Address for correspondence: Li Xu, Ohio University, Athens, OH 45701, USA. E-mail: xul@ohio.edu

Received February 19, 2020; accepted June 29, 2020.

REFERENCES

Bench, J., & Barnford, J. (1979). *Speech-Hearing Tests and the Spoken Language of Hearing Impaired Children*. Academic Press.

Chan, J. C., Freed, D. J., Vermiglio, A. J., Soli, S. D. (2008). Evaluation of binaural functions in bilateral cochlear implant users. *Int J Audiol, 47*, 296–310.

Chen, F., & Lau, A. (2014). Effect of vocoder type to Mandarin speech recognition in cochlear implant simulation. *Proc. 9th Int. Symp. Chin. Spoken Language Process.*, pp. 551–554.

Chen, F., Zheng, D., Tsao, Y. (2017). Effects of noise suppression and envelope dynamic range compression on the intelligibility of vocoded sentences for a tonal language. *J Acoust Soc Am, 142*, 1157.

Chen, Y., & Wong, L. L. N. (2017). Speech perception in Mandarin-speaking children with cochlear implants: A systematic review. *Int J Audiol, 56*(Suppl 2), S7–S16.

Darouie, A., Zamiri Abdollahi, F., Joulaie, M., Nik Nezhad, S., Ahmadi, T., Soli, S. (2020). Development of the Farsi Hearing in Noise Test. *Int J Audiol, 59*, 148–152.

Dorman, M. F., Loizou, P. C., Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J Acoust Soc Am, 102*, 2403–2411.

Dorman, M. F., Spahr, A. J., Loizou, P. C., Dana, C. J., Schmidt, J. S. (2005). Acoustic simulations of combined electric and acoustic hearing (EAS). *Ear Hear, 26*, 371–380.

Dorman, M. F., & Gifford, R. H. (2017). Speech understanding in complex listening environments by listeners fit with cochlear implants. *J Speech Lang Hear Res, 60*, 3019–3026.

Dorman, M. F., Gifford, R. H., Spahr, A. J., McKarns, S. A. (2008). The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. *Audiol Neurootol, 13*, 105–112.

Dorman, M. F., & Gifford, R. H. (2010). Combining acoustic and electric stimulation in the service of speech recognition. *Int J Audiol, 49*, 912–919.

Everhardt, M.K., Sarampalis, A., Coler, M., Başkent, D., Lowie, W. (2020). Meta-analysis on the identification of linguistic and emotional prosody in cochlear implant users and vocoder simulations. *Ear Hear*, https://doi.org/10.1097/aud.0000000000000863.

Firszt, J. B., Holden, L. K., Reeder, R. M., Skinner, M. W. (2009). Speech recognition in cochlear implant recipients: comparison of standard HiRes and HiRes 120 sound processing. *Otol Neurotol, 30*, 146–152.

Friesen, L. M., Shannon, R. V., Baskent, D., Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am, 110*, 1150–1163.

Fu, Q. J., Chinchilla, S., Galvin, J. J. (2004). The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. *J Assoc Res Otolaryngol, 5*, 253–260.

Fu, Q. J., Galvin, J. J. 3rd, Wang, X. (2017). Effect of carrier bandwidth on integration of simulations of acoustic and electric hearing within or across ears. *J Acoust Soc Am, 142*, EL561–EL566.

Gifford, R. H., Driscoll, C. L., Davis, T. J., Fiebig, P., Micco, A., Dorman, M. F. (2015). A within-subject comparison of bimodal hearing, bilateral cochlear implantation, and bilateral cochlear implantation with bilateral hearing preservation: high-performing patients. *Otol Neurotol, 36*, 1331–1337.

Gifford, R. H., Shallop, J. K., Peterson, A. M. (2008). Speech recognition materials and ceiling effects: Considerations for cochlear implant programs. *Audiol Neurootol, 13*, 193–205.

Giguère, C., Laroche, C., Soli, S. D., Vaillancourt, V. (2008). Functionally-based screening criteria for hearing-critical jobs based on the Hearing in Noise Test. *Int J Audiol, 47*, 319–328.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species–29 years later. *J Acoust Soc Am, 87*, 2592–2605.

Han, D., Liu, B., Zhou, N., Chen, X., Kong, Y., Liu, H., Zheng, Y., Xu, L. (2009). Lexical tone perception with HiResolution and HiResolution 120 sound-processing strategies in pediatric Mandarin-speaking cochlear implant users. *Ear Hear, 30*, 169–177.

Kim, B. J., Chang, S. A., Yang, J., Oh, S. H., Xu, L. (2015). Relative contributions of spectral and temporal cues to Korean phoneme recognition. *PLoS One, 10*, e0131807.

Li, J. N., Chen, S., Zhai, L., Han, D. Y., Eshraghi, A. A., Feng, Y., Yang, S. M., Liu, X. Z. (2017). The Advances in Hearing Rehabilitation and Cochlear Implants in China. *Ear Hear, 38*, 647–652.

Liang, Q., & Mason, B. (2013). Enter the dragon–China's journey to the hearing world. *Cochlear Implants Int, 14(Suppl 1)*, S26–S31.

Liu, H., Peng, X., Zhao, Y., Ni, X. (2017). The effectiveness of sound-processing strategies on tonal language cochlear implant users: A systematic review. *Pediat Inv*, *1*, 32–39.

Liu, J., & Yang, S.-M. (2019) Development of cochlear implantation and its related technique. *Chin J Otolaryngolog Skull Base Surg*, *25*, 449–455.

Loizou, P. C. (2006). Speech processing in vocoder-centric cochlear implants. *Adv Otorhinolaryngol, 64*, 109–143.

Luxford, W. M., Ad Hoc Subcommittee of the Committee on Hearing and Equilibrium of the American Academy of Otolaryngology-Head and Neck Surgery. (2001). Minimum speech test battery for postlingually deafened adult cochlear implant patients. *Otolaryngol Head Neck Surg, 124*, 125–126.

Massa, S. T., & Ruckenstein, M. J. (2014). Comparing the performance plateau in adult cochlear implant patients using HINT and AzBio. *Otol Neurotol, 35*, 598–604.

Mao, Y., & Xu, L. (2017). Lexical tone recognition in noise in normal-hearing children and prelingually deafened children with cochlear implants. *Int J Audiol, 56*(Suppl 2), S23–S30.

Meng, Q., Wang, X., Cai, Y., Kong, F., Buck, A. N., Yu, G., Zheng, N., Schnupp, J. W. H. (2019). Time-compression thresholds for Mandarin sentences in normal-hearing and cochlear implant listeners. *Hear Res, 374*, 58–68.

MSTB. The new Minimum Speech Test Battery. (2011) http://www.auditorypotential.com/MSTBfiles/MSTBManual2011-06-20%20.pdf.

Myhrum, M., Tvete, O. E., Heldahl, M. G., Moen, I., Soli, S. D. (2016). The Norwegian Hearing in Noise Test for Children. *Ear Hear, 37*, 80–92.

Nilsson, M., Soli, S. D., Sullivan, J. A. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am, 95*, 1085–1099.

Oliver, J.C., & Gonzalez, J. (2005). Gender and speaker identification as a function of the number of channels in spectrally reduced speech. *J Acoust Soc Am*, *118*, 461–470.

Qin, M. K., & Oxenham, A. J. (2006). Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *J Acoust Soc Am, 119*, 2417–2426.

Rebsher, S., Hetherington, A., Bonham, B., et al. (2018). Development and clinical introduction of the nurotron cochlear implant electrode array. *J Int Adv Otol*, *14*, 392–400.

Sargent, E. W., Herrmann, B., Hollenbeak, C. S., Bankaitis, A. E. (2001). The Minimum Speech Test Battery in profound unilateral hearing loss. *Otol Neurotol, 22*, 480–486.

Schafer, E. C., Pogue, J., Milrany, T. (2012). List equivalency of the AzBio sentence test in noise for listeners with normal-hearing sensitivity or cochlear implants. *J Am Acad Audiol, 23*, 501–509.

Shannon, R.V., Fu, Q.-J., Galvin, J. (2004). The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol*, (Suppl 552), 50–54.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303–304.

Shiroma, M., Iwaki, T., Kubo, T., Soli, S. (2008). The Japanese hearing in noise test. *Int J Audiol, 47*, 381–382.

Skinner, M. W., Arndt, P. L., Staller, S. J. (2002). Nucleus 24 advanced encoder conversion study: Performance versus preference. *Ear Hear*, *23*, 2S–16S.

Soli, S. D., Vermiglio, A., Wen, K., Filesari, C. A. (2002). Development of the hearing in noise test (HINT) in Spanish. *J Acoust Soc Am*, *112*, 2384–2384.

Soli, S. D., & Wong, L. L. (2008). Assessment of speech intelligibility in noise with the hearing in noise test. *Int J Audiol, 47*, 356–361.

Spahr, A. J., & Dorman, M. F. (2004). Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices. *Arch Otolaryngol Head Neck Surg, 130*, 624–628.

Spahr, A. J., Dorman, M. F., Loiselle, L. H. (2007). Performance of patients using different cochlear implant systems: effects of input dynamic range. *Ear Hear, 28*, 260–275.

Spahr, A. J., Dorman, M. F., Litvak, L. M., Van Wie, S., Gifford, R. H., Loizou, P. C., Loiselle, L. M., Oakes, T., Cook, S. (2012). Development and validation of the AzBio sentence lists. *Ear Hear, 33*, 112–117.

Souza, P., & Rosen, S. (2009). Effects of envelope bandwidth on the intelligibility of sine- and noise-vocoded speech. *J Acoust Soc Am, 126*, 792–805.

Su, Q., Galvin, J. J., Zhang, G., Li, Y., Fu, Q.-J. (2016) Effects of within-talker variability on speech intelligibility in Mandarin-speaking adult and pediatric cochlear implant patients. *Trends Hear*, *20*, 1–16.

Wang, W., Zhou, N., Xu, L. (2011). Musical pitch and lexical tone perception with cochlear implants. *Int J Audiol, 50*, 270–278.

Warton, D. I., & Hui, F. K. (2011). The arcsine is asinine: the analysis of proportions in ecology. *Ecology, 92*, 3–10.

Whitmal, N. A., Poissant, S. F., Freyman, R. L., Helfer, K. S. (2007). Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience. *J Acoust Soc Am, 122*, 2376–2388.

Wilson (2019) The remarkable cochlear implant and possibilities for the next large step forward. *Acoust Today*, *15*, 53–61.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature, 352*, 236–238.

Wong, L. L., & Soli, S. D. (2005). Development of the Cantonese Hearing in Noise Test (CHINT). *Ear Hear, 26*, 276–289.

Wong, L. L., Soli, S. D., Liu, S., Han, N., Huang, M. W. (2007). Development of the Mandarin Hearing in Noise Test (MHINT). *Ear Hear, 28*(2 Suppl), 70S–74S.

Xi, X., Wang, Y., Qiu, X.-Y., Shi, Y., Gao, R. (2015) Development of Mandarin AzBio sentence lists and list equivalency evaluation. *The 10th Asia Pacific Symposium on Cochlear Implant and Related Sciences*. Beijing, China.

Xu, K., Soli, S. D., Zheng, Y., Liu, S., Li, G., Tao, Y., Meng, Z. (2015). Quantification of the effects of Mandarin dialect differences on the use of norm-referenced speech perception tests. *Int J Audiol, 54*, 461–466.

Xu, L. (2016). Temporal envelopes in sine-wave speech recognition. *Interspeech 2016*, San Francisco, CA, 1682–1686.

Xu, L., & Pfingst, B. E. (2008). Spectral and temporal cues for speech recognition: Implications for auditory prostheses. *Hear Res, 242*, 132–140.

Xu, L., Thompson, C. S., Pfingst, B. E. (2005). Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am, 117*, 3255–3267.

Xu, L., Tsai, Y., Pfingst, B. E. (2002). Features of stimulation affecting tonal-speech perception: Implications for cochlear prostheses. *J Acoust Soc Am, 112*, 247–258.

Xu, L., & Zheng, Y. (2007). Spectral and temporal cues for phoneme recognition in noise. *J Acoust Soc Am, 122*, 1758.

Xu, L., & Zhou, N. (2011). Tonal languages and cochlear implants. In Zeng, F.-G., Popper, A. N., & Fay, R. R., eds. *Auditory Prostheses: New Horizons* (pp. 341–364). Springer Science+Business Media, LLC.

Yang, S. M, Li, J. N, Han, D. Y. (2010). Promotion of cochlear implant in post-lingual deafness adults in China. *Chin Arch Otolaryngol Head Neck Surg*, *17*, 23–27.

Zeng, F. G., Rebscher, S. J., Fu, Q. J., Chen, H., Sun, X., Yin, L., Ping, L., Feng, H., Yang, S., Gong, S., Yang, B., Kang, H. Y., Gao, N., Chi, F. (2015). Development and evaluation of the Nurotron 26-electrode cochlear implant system. *Hear Res, 322*, 188–199.

Zhang, N., Liu, S., Xu, J., Liu, B., Qi, B., Yang, Y., Kong, Y., Han, D. (2010). Development and applications of alternative methods of segmentation for Mandarin Hearing in Noise Test in normal-hearing listeners and cochlear implant users. *Acta Otolaryngol, 130*, 831–837.

Zhang, T., Dorman, M. F., Spahr, A. J. (2010). Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. *Ear Hear, 31*, 63–69.

Zhang, T., Spahr, A. J., Dorman, M. F., Saoji, A. (2013). Relationship between auditory function of nonimplanted ears and bimodal benefit. *Ear Hear, 34*, 133–141.

Zhou, N., Huang, J., Chen, X., Xu, L. (2013). Relationship between tone perception and production in prelingually deafened children with cochlear implants. *Otol Neurotol, 34*, 499–506.