JSLHR

Research Article

Tonal Language Speakers Are Better Able to Segregate Competing Speech According to Talker Sex Differences

Juan Zhang,[a] Xing Wang,[a] Ning-yu Wang,[a] Xin Fu,[a] Tian Gan,[a] John J. Galvin III,[b] Shelby Willis,[c] Kevin Xu,[c] Mathew Thomas,[c] and Qian-Jie Fu[c]

Purpose: The aim of this study was to compare release from masking (RM) between Mandarin-speaking and English-speaking listeners with normal hearing for competing speech when target–masker sex cues, spatial cues, or both were available.
Method: Speech recognition thresholds (SRTs) for competing speech were measured in 21 Mandarin-speaking and 15 English-speaking adults with normal hearing using a modified coordinate response measure task. SRTs were measured for target sentences produced by a male talker in the presence of two masker talkers (different male talkers or female talkers). The target sentence was always presented directly in front of the listener, and the maskers were either colocated with the target or were spatially separated from the target (+90°, –90°). Stimuli were presented via headphones and were virtually spatialized using head-related transfer functions. Three masker conditions were used to measure RM relative to the baseline condition: (a) talker sex cues, (b) spatial cues, or (c) combined talker sex and spatial cues.
Results: The results showed large amounts of RM according to talker sex and/or spatial cues. There was no significant difference in SRTs between Chinese and English listeners for the baseline condition, where no talker sex or spatial cues were available. Furthermore, there was no significant difference in RM between Chinese and English listeners when spatial cues were available. However, RM was significantly larger for Chinese listeners when talker sex cues or combined talker sex and spatial cues were available.
Conclusion: Listeners who speak a tonal language such as Mandarin Chinese may be able to take greater advantage of talker sex cues than listeners who do not speak a tonal language.

When listening to speech in competing speech background, recognition of the target sentence is most difficult when the maskers are colocated with the target, the masker talkers are the same sex as the target, and the masker speech is intelligible (e.g., Kidd et al., 2016). Both energetic masking and informational masking may attribute to this difficulty (e.g., Durlach et al., 2003). Energetic masking occurs at the periphery, where recognition of the target depends on the degree of temporal and spectral overlap with the masker. Informational masking occurs more centrally, where the target and masker are clearly audible but cannot be segregated.

Effects of Target–Masker Sex Differences on Masking Release

Differences in sex between target and the masker talkers (e.g., male target and female masker or vice versa) can allow for a large release from masking (RM) according to the acoustic properties of the competing talkers (Başkent & Gaudrain 2016; Brokx & Nooteboom 1982; Brown et al., 2010; Brungart, 2001; Brungart et al., 2001; Cullington & Zeng, 2008; Darwin et al., 2003; Darwin & Hukin, 2000; Drullman & Bronkhorst, 2004; El Boghdady et al., 2019; Vestergaard et al., 2009). The benefit from target–masker sex difference is thought to be due primarily to a reduction in informational masking (Kidd et al., 2016) and is often commonly termed "voice gender release from masking" (Oh et al., 2019). However, given the fluidity in

[a]Department of Otolaryngology, Head and Neck Surgery, Beijing Chaoyang Hospital, Capital Medical University, China
[b]House Ear Institute, Los Angeles, CA
[c]Department of Head and Neck Surgery, David Geffen School of Medicine, University of California, Los Angeles
Correspondence to Qian-Jie Fu (primary): QFu@mednet.ucla.edu; Ning-yu Wang (secondary): 2460331882@qq.com
Editor-in-Chief: Frederick (Erick) Gallun
Editor: Yi Shen
Received December 29, 2019
Revision received April 1, 2020
Accepted May 15, 2020
https://doi.org/10.1044/2020_JSLHR-19-00421

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

gender identification, it is perhaps more appropriate to use the term "talker sex release from masking" (TSRM), which we will use in this article. Previous studies have reported TSRM of approximately 10 dB with one masker talker (Cullington & Zeng, 2008) and approximately 9 dB with two masker talkers (Brown et al., 2010; Cullington & Zeng, 2008).

To segregate competing talkers when there are no spatial cues, listeners mainly rely on differences in fundamental frequency (F0; or voice pitch) and vocal tract length (VTL). Previous studies investigated the role of F0 and VTL on perception of a talker's sex (e.g., Bishop & Keating, 2012; Hillenbrand & Clark, 2009; Poon & Ng, 2015). The data from these studies showed that F0 was the primary cue for perceiving talker sex. Darwin et al. (2003) measured the effects of F0 and VTL differences on segregating target and masker speech. They found that, while differences in F0 provided systematic improvements in segregation, F0 cues alone could not account for RM according to target–masker sex differences. Similarly, differences in VTL also provided systematic improvements in segregation but could not fully account for RM according to target–masker sex differences. Interestingly, the authors also found that differences in intonation across talkers could play as large a role as F0 cues in determining overall segregation performance.

Besides its importance as a cue to segregate competing talkers, F0 is also the primary cue for recognizing tones in tonal languages, such as Mandarin Chinese, where lexical tones convey linguistic meaning (Liang, 1963). Xie and Myers (2015) found that lexical tone experience improved accuracy in talker identification, possibly because tonal language experience generally improved pitch perception acuity. Such enhanced pitch perception via tonal language experience may also improve segregation of competing talkers, especially when target–masker sex cues are available.

Recently, Chen et al. (2020) measured speech recognition thresholds (SRTs) in Mandarin-speaking listeners with normal hearing (NH) with one, two, or four masker talkers, using a combination of target–masker vocal characteristics, similar to that of Cullington and Zeng (2008). Chen et al. reported a TSRM of 12.3 dB with a one-talker masker and 11.7 dB with a two-talker masker. This appears to be somewhat larger than the TSRM reported for English-speaking listeners with NH in the study of Cullington and Zeng (10.2 and 8.4 dB for the one- and two-talker maskers, respectively). Due to the potential role of intonation cues on talker segregation (Darwin et al., 2003), the tonal patterns of individual Chinese tones may have contributed to the performance difference observed between English and Chinese listeners. In addition to language differences (tonal vs. nontonal), the discrepancy observed in TSRM may have also been due to other factors such as testing materials and testing protocols. In the study of Chen et al., the testing materials consisted of matrix-style sentences (Tao et al., 2017, 2018) for both the target and maskers. In the study of Cullington and Zeng, Hearing in Noise Test sentences (Nilsson et al., 1994) were used as the target and IEEE sentences (Rothauser et al., 1969) were used for the maskers. Also, a closed-set testing paradigm was used by Chen et al., while an open-set testing paradigm was used by Cullington and Zeng.

Effects of Spatial Separation on Masking Release

Spatial cues have been shown to improve recognition of target speech in the presence of competing maskers (e.g., Bronkhorst & Plomp, 1988; Brown et al., 2010; Dirks & Wilson, 1969; Freyman et al., 1999, 2001; Hawley et al., 1999; Hu et al., 2018; Kidd et al., 1998, 2016; Noble & Perrett, 2002). Spatial RM (SRM) is generally defined as the improvement in target speech understanding improved when targets and maskers are spatially separated. SRM is primarily due to a reduction in informational masking (Kidd et al., 2016) and may be partly attributed to head shadow and binaural processing effects (e.g., Akeroyd, 2006; Culling et al., 2004; Hawley et al., 2004).

The amount of SRM may be affected by many factors, including but not limited to stimulus type (e.g., Brown et al., 2010; Hu et al., 2018; Kidd et al., 2010), hearing status (e.g., Ching et al., 2011; Glyde et al., 2015; Srinivasan et al., 2016, 2017), and listener age (e.g., Besser et al., 2015; Brown et al., 2010; Johnstone & Litovsky, 2006; Srinivasan et al., 2016, 2017; Yuen & Yuan, 2014; Zobel et al., 2019). Typically, speech maskers produce larger SRM than do steady speech-shaped noise maskers (e.g., Freyman et al., 1999; Hu et al., 2018). Listeners with NH typically have larger SRM than do listeners with hearing impairment (e.g., Hu et al., 2018; Srinivasan et al., 2016, 2017). Young adults typically have larger SRM than do elderly adults (e.g., Besser et al., 2015; Srinivasan et al., 2016, 2017) or children (e.g., Brown et al., 2010; Yuen & Yuan, 2014). Jakien et al. (2017) also found that SRM for all listeners, regardless of age or hearing status, improved with increasing overall sensation level and/or bandwidth. However, the magnitude of these level and bandwidth effects was small compared to the general benefit of spatial cues. SRM is relatively robust in adults with NH when the two maskers are symmetrically placed relative to the target (e.g., Brown et al., 2010; Hu et al., 2018; Kidd et al., 2010). Previous studies reported an SRM of approximately 12 dB for the recognition of a target sentence presented at 0° in the presence of two speech maskers placed symmetrically at ± 90° (e.g., 11.6 dB in Marrone et al., 2008; 12.1 dB in Kidd et al., 2010; 12.0 dB in Brown et al., 2010).

Besides the contribution of F0 cues to TSRM (Darwin et al., 2003), intonation cues embedded in Mandarin Chinese tones may also contribute to degree of SRM. Wu et al. (2005) investigated the effect of perceived spatial separation induced by the precedence effect on SRM in Mandarin-speaking listeners with NH. They reported an SRM of 3.3 dB for the Mandarin-speaking listeners, somewhat smaller than the 4–9 dB of SRM reported by Freyman et al. (1999) for English-speaking listeners with NH using a similar test paradigm. Wu et al. suggested that, while perceived spatial cues may reduce perceived target–masker spatial similarity, they may also interact with other dimensions of

target–masker similarities, which may differ across tonal and nontonal languages. In a follow-up study, Wu et al. measured SRM in both Chinese and English listeners with NH when masker sentences were either the same language (i.e., Chinese target with Chinese maskers, English target with English maskers) or different languages (i.e., Chinese target with English maskers, English target with Chinese maskers). Both listener groups benefitted equally from perceived spatial cues when the target–masker language was different. However, Chinese listeners benefitted less from perceived spatial cues than did English listeners when the target–masker language was the same.

Effects of Talker Sex Cues and Spatial Cues on Masking Release

Most previous studies have focused on the independent effects of target sex cues and spatial cues on RM. Although some studies have examined the combined effects of talker and spatial cues on RM (Rennies et al., 2019), only a few studies have evaluated the combined effects of talker sex and spatial cues on RM using the Listening in Spatialized Noise–Sentences (LiSN-S) Test. The LiSN-S Test is a validated clinical test with which to assess auditory stream segregation by creating a three-dimensional auditory environment for stimuli delivered via headphones (Brown et al., 2010; Cameron & Dillon, 2007, 2008; Cameron et al., 2009, 2011). The LiSN-S Test can be used to evaluate RM according to talker sex cues (TSRM), spatial cues (SRM), or both talker sex and spatial cues (TSSRM). Brown et al. (2010) collected normative data and test–retest reliability data for adolescents and young adults with NH using the LiSN-S Test. The amount of TSRM, SRM, and TSSRM increased as a function of age at testing. For young adults, the mean improvement in SRT was approximately 9, 12, and 14 dB with talk sex cues, spatial cues, or combined talker sex and spatial cues, respectively, suggesting that talker sex and spatial cues provide greater RM than either cue alone.

Rationale of This Study

The data from the studies of Cullington and Zeng (2008) and Chen et al. (2020) suggest that Chinese-speaking listeners may benefit more from talker sex cues than do English-speaking listeners; however, differences in speech materials and test paradigms make direct comparison difficult. Differences between tonal (Chinese) and nontonal (English) languages may partly contribute to differences in talker sex cue utilization. Tone patterns generally include much larger F0 variations in Chinese than in English. Deroche et al. (2019) found that Chinese listeners had better sensitivity to dynamic changes in F0 than did English listeners, especially for large F0 variations. As noted above, intonation cues in Mandarin Chinese (Darwin et al., 2003) and/or improved pitch acuity associated with long-term tonal language experience (Xie & Myers, 2015) may provide some advantage for Chinese listeners. Different from utilization of talker sex cues, Wu et al. (2005, 2011) found that Chinese listeners may benefit less from spatial cues than do English listeners. Thus, Chinese listeners may benefit more from talker sex cues but less from spatial cues than do English listeners. However, the effects of language on TSSRM are unclear and warrant further investigation.

The goal of this study was to determine the effects of language on the amount of RM when talker sex cues and/or spatial cues are available. To minimize the effects of testing materials and testing protocols, SRTs were measured using a similar approach, as in the LiSN-S Test (e.g., Brown et al., 2010), except that five-word matrix-style sentences and a closed-set coordinate response measure test paradigm were used. Chinese listeners were tested while listening to Chinese target and masker speech, and English listeners were tested while listening to English target and masker speech; for both groups of listeners, the stimuli were similar, and the test paradigm was identical. Consistent with previous studies, we predicted that Chinese listeners would experience greater RM from talker sex cues but less RM from spatial cues than would English listeners. While we expected RM to be greatest when both cues were available to both listening groups, it was unclear how cue utilization might differ between groups when both talker sex and spatial cues were available

Method

Subjects

Twenty-one Mandarin-speaking Chinese adults with NH (10 men and 11 women; mean age at testing = 23.2 years, range: 21–33 years) and 15 English-speaking adults with NH (6 men and 9 women; mean age at testing = 24.9 years, range: 20–41 years) participated in the study. All participants had pure-tone thresholds of < 20 dB HL at all audiometric frequencies between 250 and 8000 Hz. In compliance with ethical standards for human subjects, written informed consent was obtained from all participants before proceeding with any of the study procedures. This study was approved by the institutional review board in Chaoyang Hospital, Capital Medical University (Chinese listeners), and the University of California, Los Angeles (English listeners).

Test Materials

The matrix-style test materials were drawn from the Closed-Set Mandarin Speech corpus (Tao et al., 2017) for Chinese listeners and from the Sung Speech Corpus (Crew et al., 2015, 2016) for English listeners. In both cases, target and masker stimuli consisted of five-word sentences, designed according to matrix-style test paradigms. To create the target sentences, one of the 10 words was chosen at random from each of five categories (Name, Verb, Number, Color, and Object); for both Chinese and English, the target speech was produced by a single male talker. Similarly, masker sentences were created by choosing one of the 10 words from each category that was not used for the target sentence; masker sentences each contained unique

words. Masker sentences were produced by two male talkers who were different from the target male talker or by two female talkers. F0 was averaged across all stimuli for each Chinese and English talker. Table 1 shows the 10%, 50%, and 90% quantiles of F0s estimated for each Chinese and English talker. The size of F0 excursion was estimated in terms of the semitone difference between the 10% and 90% quantiles.

In this study, SRTs were defined as the target-to-masker ratio (TMR) that produced 50% correct recognition of target keywords in sentences. SRTs were adaptively measured (e.g., Brungart et al., 2001; Tao et al., 2017, 2018). Two target keywords (randomly selected from the Number and Color categories) were embedded in a five-word carrier sentence uttered by the male target talker. For Mandarin Chinese, the first word in the target sentence was always the Name "**Xiaowang**," followed by randomly selected words from the other remaining categories. For example, the target sentence could be (in Chinese) "**Xiaowang** sold ***Three Red*** strawberries" or "**Xiaowang** chose ***Four Brown*** bananas" and so forth. (Name to cue the target talker in **bold**; keywords in ***bold italics***). For English, the first word in the target sentence was always the Name "**John**," followed by randomly generated words from the remaining categories. Thus, the target sentence could be "**John** moves ***Six Gold*** pants" or "**John** needs ***Two Green*** shoes" and so forth.

Similar to the target sentence, two masker sentences were also generated for each test trial; words were randomly selected from each category but excluded the words used in the target sentence and the other masker sentence. Thus, the target and masker sentences all contained different words in each category. An example Chinese target sentence could be "**Xiaowang** sold ***Three Red*** strawberries," while the masker sentences could be the combination "Xiaozhang saw *Two Blue* kumquats" and "Xiaodeng took *Eight Green* papayas." An example of an English target sentence could be "**John** moves ***Six Gold*** pants," while the masker sentences could be "Bob Finds *Two Blue* coats" and "Greg loans *Five Grey* Jeans."

Listening Conditions

All target and masker stimuli were presented via Sennheiser HDA200 headphones connected to the output of an audio interface (Edirol UA-25) connected to a PC. Nonindividualized head-related transfer functions (HRTFs) were used to create a virtual auditory space for headphone presentation (Wightman & Kistler, 1989). The target sentence originated directly in front of the listener (0°), and the two masker sentences were either colocated with the target (0°) or presented to the left (−90°) and right of the target (+90°) separately. The masker talker sex was either the same as or different from the male target talker. Accordingly, there were four listening conditions: (a) baseline (no talker sex or spatial cues), (b) talker sex cues (TS condition), (c) spatial cues (S condition), and (d) combined talker sex and spatial cues (TSS condition).

Test Protocols

The target sentence was always presented at 65 dBA, while the level of masker sentences was globally adjusted according to the correctness of the listener's response; note that the root-mean-square amplitude of the masker sentences was normalized before adjustment. For example, for a TMR of +10 dB, the level of the target sentence was 65 dBA and the level of each masker sentence was 55 dBA. During each test trial, sentences were presented at the target TMR; the initial TMR was 10 dB for the baseline condition (no talker sex or spatial cues) and 0 dB for the TS, S, and TSS conditions. Participants were instructed to listen to the target sentence (produced by the male target talker and beginning with the name "**Xiaowang**" for Chinese listeners and "**John**" for English listeners) and then click on one of the 10 response choices from each of the Number and Color categories; no selections could be made from the remaining categories, which were grayed out. If the participant identified both key words correctly, the TMR was reduced; if the participant did not identify both key words, the TMR was increased. The initial step size was 4 dB, and

Table 1. Distribution of fundamental frequency (F0) estimated across all stimuli for each Chinese and English talker.

Talker	10% Quantile (Hz)	50% Quantile (Hz)	90% Quantile (Hz)	F0 excursion size (semitones)
Chinese				
Target (male)	80.8	124.0	198.6	15.6
Male 1	89.9	128.5	207.6	14.5
Male 2	89.5	156.3	279.7	19.7
Female 1	133.5	177.2	318.9	15.1
Female 2	152.3	225.9	357.1	14.8
English				
Target (male)	89.7	103.5	116.6	4.5
Male 1	113.2	124.1	133.8	2.9
Male 2	77.5	88.7	100.1	4.4
Female 1	161.0	173.0	188.9	2.8
Female 2	132.7	157.3	185.1	5.8

Note. F0 excursion size was estimated as the semitone difference between the 10% and 90% quantiles.

the final step size was 2 dB. The SRT was calculated by averaging the last six reversals in TMR. If there were fewer than six reversals within 20 trials, the test run was discarded and another run was measured. Three test runs were completed for each listening condition, and the SRT was averaged across runs. The listening conditions and repetitions were randomized within and across participants.

Results

Figure 1 shows the box plots of SRTs for the four listening conditions in Chinese and English participants. For Chinese listeners, mean SRTs were −0.02 ± 0.46, −11.67 ± 2.96, −13.73 ± 2.01, and −18.13 ± 2.47 dB for the baseline, TS, S, and TSS conditions, respectively. For English listeners, mean SRTs were 0.36 ± 1.95, −8.31 ± 1.81, −11.87 ± 2.79, and −12.62 ± 2.69 dB for the baseline, TS, S, and TSS conditions, respectively. A repeated-measures analysis of variance was performed on the data shown in Figure 1, with listening condition (baseline, TS, S, and TSS) as the within-subject factor and language (Chinese, English) as the between-subjects factor. Results showed significant effects of listening condition, $F(3, 102) = 492.3$, $p < .001$, $\eta_p^2 = .935$, and language, $F(1, 34) = 23.3$, $p < .001$, $\eta_p^2 = .407$, on SRTs; there was a significant interaction, $F(3, 102) = 12.7$, $p < .001$, $\eta_p^2 = .272$. Post hoc Bonferroni pairwise comparisons showed that Chinese listeners performed significantly better in the TS ($p < .001$), S ($p = .026$), and TSS ($p < .001$) listening conditions. However, there was no significant difference between Chinese and English listeners in the baseline condition ($p = .506$). Bonferroni pairwise comparisons also revealed significant differences among all listening conditions in Chinese listeners ($p < .005$ in all cases). However, for English listeners, SRTs were significantly different across all listening conditions ($p < .05$ in all cases), except for between S and TSS ($p > .999$).

The amount of RM was calculated as the reduction in SRTs with the TS, S, and TSS listening conditions, relative to baseline. Figure 2 shows the box plots of RM for the TS, S, and TSS listening conditions, for Chinese and English listeners. For Chinese listeners, mean RM was 11.65 ± 2.73, 13.71 ± 2.7, and 18.11 ± 2.80 dB for the TS, S, and TSS conditions, respectively. For English listeners, mean RM was 8.67 ± 1.82, 12.23 ± 2.77, and 12.98 ± 2.55 dB for the TS, S, and TSS conditions, respectively. A repeated-measures analysis of variance was performed on the data shown in Figure 2, with RM type (TS, S, and TSS) as the within-subject factor and language (Chinese, English) as the between-subjects factor. Results showed significant effects of RM type, $F(2, 68) = 80.4$, $p < .001$, $\eta_p^2 = .703$, and language group, $F(1, 34) = 18.7$, $p < .001$, $\eta_p^2 = .355$; there was also a significant interaction, $F(2, 68) = 9.3$, $p < .001$, $\eta_p^2 = .216$. Post hoc Bonferroni pairwise comparisons showed that Chinese listeners exhibited significantly larger RM for the TS ($p = .001$) and TSS ($p < .001$) conditions than did English listeners. However, no significant difference between Chinese and English listeners was observed for the S condition ($p = .123$). Bonferroni pairwise comparisons also showed that, for Chinese listeners, RM was significantly larger for the TSS than for the TS and S conditions ($p < .001$ in both cases) and that RM was significantly

Figure 1. Box plots of speech recognition threshold (SRTs; in dB TMR) as a function of listening condition for Chinese and English listeners. The boxes show the 25th and 75th percentiles, the horizontal solid line shows the median, the dashed line shows the mean, the error bars show the 10th and 90th percentiles, and the circles display outliers. The baseline condition contained no talker sex or spatial cues. The TS condition contained talker sex cues, but no spatial cues. The S condition contained spatial cues, but no talker sex cues. The TSS condition contained both talker sex and spatial cues.

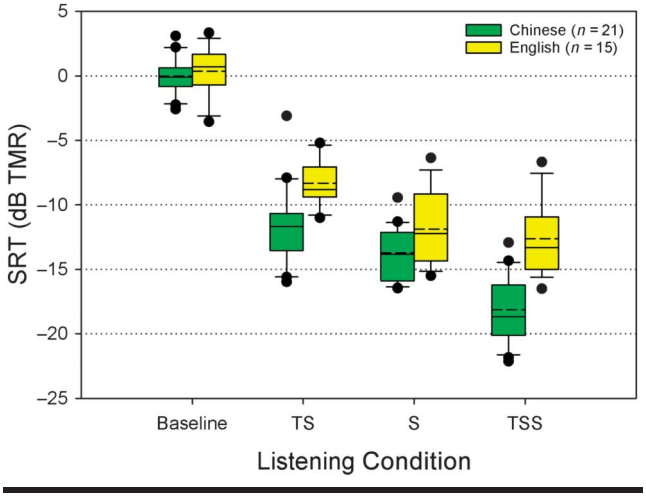
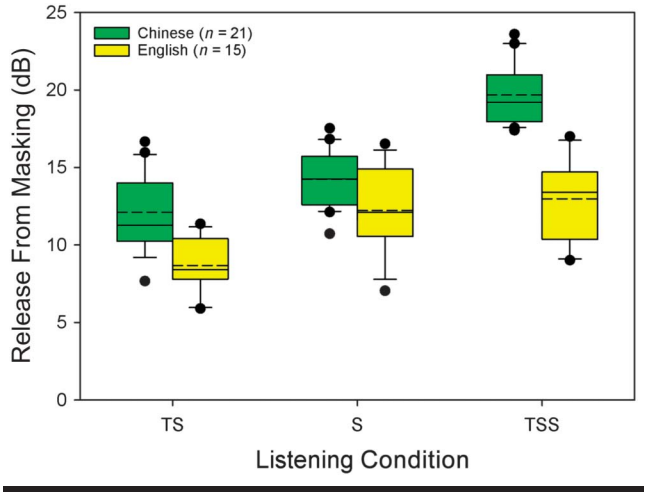


Figure 2. Box plots for release from masking (RM) for the talker sex cues (TS), spatial cues (S), and combined talker sex and spatial cues (TSS) listening conditions; RM was calculated relative to baseline. The boxes show the 25th and 75th percentiles, the horizontal solid line shows the median, the dashed line shows the mean, the error bars show the 10th and 90th percentiles, and the circles display outliers.



larger for the S than for the TS condition ($p = .002$). Similarly, for English listeners, RM was significantly larger for the TSS than for the TS and S conditions ($p < .001$ in both cases); however, there was no significant difference between the S and TSS conditions ($p = .859$).

Discussion

The data from this study demonstrated a clear benefit of talker sex and spatial cues for segregation of competing speech in both Chinese and English listeners. For English listeners, the amount of RM in this study was consistent with that reported for the similar LiSN-S Test in Brown et al. (2010). Chinese and English listeners exhibited similar RM when spatial cues were available, but Chinese listeners exhibited significantly larger RM when talker sex cues or combined talker sex and spatial cues were available. This suggests that talker sex cues (with or without spatial cues) facilitate the segregation of competing speech to a greater extent for Chinese listeners than for English listeners.

Effects of Language on Baseline Speech Performance

When listening to competing speech, recognition of the target sentence is most difficult when masker sentences are colocated with the target, masker talkers are the same sex as the target, and the masker speech is intelligible. The perceptual similarity among target and masker talkers may be affected by many factors, including mean F0 and F0 variation (F0 excursion size). Darwin et al. (2003) found that differences in F0 greater than 2 semitones produced systematic improvements in segregation performance. They also found that the natural variations in intonation (F0 excursion size) in the utterances spoken by one talker were so large that introducing artificial changes to F0 or VTL of the competing phrases provided no additional improvement. In this study, the mean F0 difference between the male target and male maskers was 0.2 semitones for English and 2.3 semitones for Chinese. For English, mean F0 excursion size was similar between the male target (4.5 semitones) and male maskers (3.7 semitones). For Chinese, while the mean F0 excursion size was also similar between the male target (15.6 semitones) and male maskers (17.1 semitones), the overall F0 excursion was approximately 4 times greater than observed for English talkers. Despite the large differences between Chinese and English talkers in terms of mean F0 and F0 pitch excursion, the differences between target and masker talkers within each language were relatively small. Under these conditions, putative advantages from long-term experience with tonal language (Xie & Myers, 2015) may not be beneficial. Indeed, there was no significant difference in baseline SRTs between Chinese and English listeners. These results suggest that, without strong segregation cues, F0 differences contribute little to obligatory segregation (Moore & Gockel, 2012) for both a tonal and nontonal languages. Similarly, large F0 variations contributed little to obligatory segregation if the target and masker talkers had similar F0 excursion sizes.

The present baseline data were also consistent with previous studies. For example, Chen et al. (2020) reported an SRT of −3.1 dB (in terms of signal-to-noise ratio [SNR]) for the recognition of a target sentence in the presence of two same-sex masker sentences in Chinese listeners. Note that the SNR in Chen et al.'s study represents the ratio between the target talker and the combined two-talker masker sentences. SRTs were expressed in terms of TMR in this study to allow easier comparison between the colocated and spatially separated masker conditions; TMR represents the ratio between the target sentence and each of the masker sentences (which were normalized to have the same root-mean-square amplitude). Thus, in this study, when the levels of all three talkers were equal in the colocated condition, the TMR would be 0 dB and the SNR would be approximately −3.0 dB. Therefore, the SRT of −3.1 dB SNR reported by Chen et al. was equivalent to 0.10 dB TMR, consistent with the −0.02 dB reported for the Chinese listeners in this study. For previous studies with English listeners, SRTs for colocated competing speech ranged from 1.9 dB TMR (−1.1 dB SNR) in the study of Cullington and Zeng (2008) to −1.6 dB TMR in the study of Brown et al. (2010). The baseline SRT obtained in this study was 0.4 dB TMR, in between the SRTs of these previous studies. Interestingly, the difference in SRTs was quite small among these studies, despite differences in testing materials, testing protocols, and language.

Effects of Language on Utilization of Talker Sex Cues

The benefits of talker sex cues on RM have been well documented (e.g., Başkent & Gaudrain 2016; Brokx & Nooteboom, 1982; Brown et al., 2010; Brungart, 2001; Brungart et al., 2001; Cullington & Zeng, 2008; Darwin et al., 2003; Darwin & Hukin, 2000; Drullman & Bronkhorst, 2004; El Boghdady et al., 2019; Vestergaard et al., 2009). With two masker talkers, Cullington and Zeng (2008) and Brown et al. (2010) reported an RM of approximately 9 dB in English listeners when talker sex cues were available, close to the 8.7 dB of RM reported in this study. Chen et al. (2020) reported an RM of 11.7 dB in Chinese listeners when talker sex cues were available, identical to the 11.7 dB TSRM in this study. In this study, RM with talker sex cues was significantly better for Chinese listeners (11.7 dB) than for English listeners (8.7 dB; $p = .002$). The present data support our prediction that talker sex cues would be better utilized by Chinese listeners.

Mean F0 and mean F0 excursion size were estimated for all talkers to explore whether the difference in RM with talker sex cues between English and Chinese listeners was driven by acoustic differences. The mean F0 difference between the male target and the female maskers was 8.1 semitones for English and 8.3 semitones for Chinese. The mean F0 excursion sizes were 4.5 semitones for the male target and 4.0 semitones for all maskers (averaged from two male

and two female maskers) in English and 15.6 semitones for the male target and 16.0 semitones for all maskers in Chinese. These data suggest that F0 differences between target and masker talkers cannot solely account for the difference in RM with talker sex cues between Chinese and English listeners. As suggested by Xie and Myers (2015), long-term experience with tonal languages may have increased pitch perception acuity in Chinese listeners. Deroche et al. (2019) also found that Chinese listeners had better sensitivity to dynamic F0 cues than did English listeners, especially for large F0 variations. The overall F0 variation in the TS condition, estimated in terms of the semitone difference between the 10% quantiles of F0s for the male target talker and 90% quantiles of F0s for the two female masker talkers, was quite large (24.8 semitones) for the present Chinese listeners. Thus, better pitch perception and better sensitivity to dynamic F0 cues due to long-term experience with tonal language may have contributed to the larger RM with talker sex cues observed in Chinese listeners. However, this study cannot rule out the possibility that the large RM with talker sex cues in Chinese listeners may be partly due to other acoustic differences (e.g., VTL) between Chinese and English stimuli.

Effects of Language on Utilization of Spatial Cues

The benefits of spatial cues for RM have been well documented in the literature (e.g., Bronkhorst & Plomp, 1988; Brown et al., 2010; Dirks & Wilson, 1969; Freyman et al., 1999, 2001; Hawley et al., 1999; Hu et al., 2018; Kidd et al., 1998, 2016; Noble & Perrett, 2002). Robust SRM in English-speaking adults with NH have been reported for two symmetrically spaced speech maskers at ± 90° (e.g., 11.6 dB in Marrone et al., 2008; 12.1 dB in Kidd et al., 2010; 12.0 dB in Brown et al., 2010). In this study, SRM was 12.2 dB, similar to SRM reported in previous studies. While mean SRM was slightly better for Chinese listeners (13.7 dB), there was no significant difference in SRM between Chinese and English listeners ($p = .123$).

The present results were contrary to our prediction and previous findings that showed that Chinese listeners exhibited less SRM than did English listeners. Wu et al. (2005, 2011) suggested that, while spatial cues may reduce perceived target–masker spatial similarity, it may also interact with other dimensions of targets and maskers (e.g., target–masker sex differences), which may differ between languages. Note that Wu et al. used delays between the left and right speakers (precedence effects) to introduce spatial cues. The SRM from this manipulation was smaller than that for studies that manipulated spatial cues via physical speaker locations or via virtual speaker locations in headphones using HRTFs. This suggests that spatial cues introduced by precedence effects may not be as strong as those induced by physical or virtual spatial cues. In instances with stronger spatial cues, interactions with other dimensions of target–masker similarities may also be reduced, resulting in similar SRM between English and Chinese listeners (as observed in this study). The present data showed robust SRM for recognition of target speech in the presence of symmetrically placed maskers (± 90°), regardless of languages.

Effects of Language on Utilization of Combined Talker Sex and Spatial Cues

As noted in the introduction, most previous studies evaluated utilization of talker sex and spatial cues independently. Brown et al. (2010) studied utilization of combined talker sex and spatial cues in English listeners using the LiSN-S Test, finding that RM was better with combined talker sex and spatial cues (14 dB) than with spatial cues alone (12 dB). In this study, for English listeners, RM was slightly (but not significantly) better with combined talker sex and spatial cues (13.0 dB) than with spatial cues alone (12.2 dB). Still, RM for the present English listeners was comparable to that in Brown et al.'s study. For Chinese listeners, RM significantly increased from 13.7 dB with spatial cues alone to 18.1 dB with combined talker sex and spatial cues ($p < .001$). Such improvement is not surprising since talker sex cues may facilitate the segregation of competing speech, with or without spatial cues, given the greater sensitivity to dynamic pitch (Deroche et al., 2019) and sharpened pitch acuity (Xie & Myers, 2015) associated with long-term tonal language experience.

Clinical Implications and Future Work

The present data with English listeners showed that the amount of RM with the TS, S, and TSS cues was comparable to that with the LiSN-S Test in the study of Brown et al. (2010), despite some differences in test materials and methodology. As such, the present materials and methodology may be a useful research and clinical tool with which to measure segregation of target speech from maskers. One advantage for the current methodology is that tests can be self-administered. This may be especially helpful for monitoring progress with training programs such as LiSN & Learn (Cameron et al., 2012) or other spatial hearing training programs. The present software can also be modified to provide trial-by-trial feedback, allowing it to be used both as a testing and training tool. Indeed, such training may help listeners to better use talker sex and/or spatial cues to segregate competing speech. Training may be especially beneficial for listeners with hearing impairment (e.g., cochlear implant patients) who have limited SRM (Hu et al., 2018) and TSRM (Chen et al., 2020) due to spectral degradation.

The present data support the notion that the better RM observed in Chinese listeners when talker sex cues or combined talker sex and spatial cues were available may be driven by better sensitivity to dynamic F0 cues (Deroche et al., 2019) and sharpened pitch acuity (Xie & Myers, 2015) associated with long-term tonal language experience. Unfortunately, pitch perception was not directly measured in this study. Future studies may reveal the relationships among tonal language experience, pitch perception, and

utilization of talker sex cues for segregation of competing speech.

Summary and Conclusion

Understanding of target sentences was measured in the presence of two colocated or spatially separated masker talkers in adult Chinese and English listeners with NH. The masker talker sex was the same or different from the target male talker. Stimuli were presented via headphones using HRTFs to create virtual sound sources. SRTs were measured using a closed-set modified coordinate response measure task. Major findings include the following:

1. For both listener groups, large amounts of RM were observed when talker sex and/or spatial cues were available.
2. While there was no significant difference in baseline SRTs (no talker sex or spatial cues) between Chinese and English listeners, SRTs were significantly lower for Chinese listeners when talker sex and/or spatial cues were available.
3. While there was no significant difference in SRM between Chinese and English listeners, RM was significantly larger in Chinese listeners when talker sex cues or combined talker sex and spatial cues were available. The greater ability to utilize talker sex cues may be due to Chinese listeners' long-term experience with tonal languages, where pitch cues are lexically meaningful.

Acknowledgments

This work was supported by the National Institutes of Health (R01-DC016883, awarded to Qian-Jie Fu). We thank all of the participants for their contribution to this study.

References

Akeroyd, M. A. (2006). The psychoacoustics of binaural hearing. *International Journal of Audiology, 45*(Suppl. 1), 25–33. https://doi.org/10.1080/14992020600782626

Başkent, D., & Gaudrain, E. (2016). Musician advantage for speech-on-speech perception. *The Journal of the Acoustical Society of America, 139*(3), EL51–EL56. https://doi.org/10.1121/1.4942628

Besser, J., Festen, J. M., Goverts, S. T., Kramer, S. E., & Pichora-Fuller, M. K. (2015). Speech-in-speech listening on the LiSN-S Test by older adults with good audiograms depends on cognition and hearing acuity at high frequencies. *Ear and Hearing, 36*(1), 24–41. https://doi.org/10.1097/AUD.0000000000000096

Bishop, J., & Keating, P. (2012). Perception of pitch location within a speaker's range: Fundamental frequency, voice quality and speaker sex. *The Journal of the Acoustical Society of America, 132*(2), 1100–1112. https://doi.org/10.1121/1.4714351

Brokx, J. P. L., & Nooteboom, S. G. (1982). Intonation and the perceptual separation of simultaneous voices. *Journal of Phonetics, 10*(1), 23–36. https://doi.org/10.1016/S0095-4470(19)30909-X

Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *The Journal of the Acoustical Society of America, 83*(4), 1508–1516. https://doi.org/10.1121/1.395906

Brown, D. K., Cameron, S., Martin, J. S., Watson, C., & Dillon, H. (2010). The North American Listening in Spatialized Noise–Sentences Test (NA LiSN-S): Normative data and test–retest reliability studies for adolescents and young adults. *Journal of the American Academy of Audiology, 21*(10), 629–641. https://doi.org/10.3766/jaaa.21.10.3

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America, 109*(3), 1101–1109. http://doi.org/10.1121/1.1345696

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America, 110*(5 Pt. 1), 2527–2538. https://doi.org/10.1121/1.1408946

Cameron, S., Brown, D., Keith, R., Martin, J., Watson, C., & Dillon, H. (2009). Development of the North American Listening in Spatialized Noise–Sentences Test (NA LiSN-S): Sentence equivalence, normative data, and test–retest reliability studies. *Journal of the American Academy of Audiology, 20*(2), 128–146. https://doi.org/10.3766/jaaa.20.2.6

Cameron, S., & Dillon, H. (2007). Development of the Listening in Spatialized Noise–Sentences test (LISN-S). *Ear and Hearing, 28*(2), 196–211. https://doi.org/10.1097/AUD.0b013e318031267f

Cameron, S., & Dillon, H. (2008). The Listening in Spatialized Noise–Sentences test (LISN-S): Comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *Journal of the American Academy of Audiology, 19*(5), 377–391. https://doi.org/10.3766/jaaa.19.5.2

Cameron, S., Glyde, H., & Dillon, H. (2011). Listening in Spatialized Noise–Sentences Test (LiSN-S): Normative and retest reliability data for adolescents and adults up to 60 years of age. *Journal of the American Academy of Audiology, 22*(10), 697–709. http://doi.org/10.3766/jaaa.22.10.7

Cameron, S., Glyde, H., & Dillon, H. (2012). Efficacy of the LiSN & Learn auditory training software: Randomized blinded controlled study. *Audiology Research, 2*(1), e15. https://doi.org/10.4081/audiores.2012.e15

Chen, B., Shi, Y., Zhang, L., Sun, Z., Li, Y., Gopen, Q., & Fu, Q.-J. (2020). Masking effects in the perception of multiple simultaneous talkers in normal-hearing and cochlear implant users. *Trends in Hearing, 24*. https://doi.org/10.1177/2331216520916106

Ching, T. Y. C., van Wanrooy, E., Dillon, H., & Carter, L. (2011). Spatial release from masking in normal-hearing children and children who use hearing aids. *The Journal of the Acoustical Society of America, 129*(1), 368–375. https://doi.org/10.1121/1.3523295

Crew, J. D., Galvin, J. J., III, & Fu, Q. J. (2015). Melodic contour identification and sentence recognition using sung speech. *The Journal of the Acoustical Society of America, 138*(3), EL347–EL351. https://doi.org/10.1121/1.4929800

Crew, J. D., Galvin, J. J., III, & Fu, Q. J. (2016). Perception of sung speech in bimodal cochlear implant users. *Trends in Hearing, 20,* 2331216516669329. https://doi.org/10.1177/2331216516669329

Culling, J. F., Hawley, M. L., & Litovsky, R. Y. (2004). The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *The Journal of the Acoustical Society of America, 116*(2), 1057–1065. https://doi.org/10.1121/1.1772396

Cullington, H. E., & Zeng, F.-G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-

hearing, cochlear-implant, and implant simulation subjects. *The Journal of the Acoustical Society of America, 123*(1), 450–461. https://doi.org/10.1121/1.2805617

Darwin, C. J., Brungart, D. S., & Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *The Journal of the Acoustical Society of America, 114*(5), 2913–2922. https://doi.org/10.1121/1.1616924

Darwin, C. J., & Hukin, R. W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *The Journal of the Acoustical Society of America, 107*(2), 970–977. https://doi.org/10.1121/1.428278

Deroche, M. L. D., Lu, H.-P., Kulkarni, A. M., Caldwell, M., Barrett, K. C., Peng, S.-C., Limb, C. J., Lin, Y.-S., & Chatterjee, M. (2019). A tonal-language benefit for pitch in normally-hearing and cochlear-implanted children. *Scientific Reports, 9,* 109. https://doi.org/10.1038/s41598-018-36393-1

Dirks, D. D., & Wilson, R. H. (1969). The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech and Hearing Research, 12*(1), 5–38. https://doi.org/10.1044/jshr.1201.05

Drullman, R., & Bronkhorst, A. W. (2004). Speech perception and talker segregation: Effects of level, pitch, and tactile support with multiple simultaneous talkers. *The Journal of the Acoustical Society of America, 116*(5), 3090–3098. https://doi.org/10.1121/1.1802535

Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003). Note on informational masking (L). *The Journal of the Acoustical Society of America, 113*(6), 2984–2987. https://doi.org/10.1121/1.1570435

El Boghdady, N., Gaudrain, E., & Başkent, D. (2019). Does good perception of vocal characteristics relate to better speech-on-speech intelligibility for cochlear implant users? *The Journal of the Acoustical Society of America, 145*(1), 417–439. https://doi.org/10.1121/1.5087693

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America, 109*(5, Pt. 1), 2112–2122. https://doi.org/10.1121/1.1354984

Freyman, R. L., Helfer, K. S., Mccall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America, 106*(6), 3578–3588. https://doi.org/10.1121/1.428211

Glyde, H., Buchholz, J. M., Nielsen, L., Best, V., Dillon, H., Cameron, S., & Hickson, L. (2015). Effect of audibility on spatial release from speech-on-speech masking. *The Journal of the Acoustical Society of America, 138*(5), 3311–3319. https://doi.org/10.1121/1.4934732

Hawley, M. L., Litovsky, R. Y., & Colburn, H. S. (1999). Speech intelligibility and localization in a multi-source environment. *The Journal of the Acoustical Society of America, 105*(6), 3436–3448. https://doi.org/10.1121/1.424670

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *The Journal of the Acoustical Society of America, 115*(2), 833–843. https://doi.org/10.1121/1.1639908

Hillenbrand, J. M., & Clark, M. J. (2009). The role of f0 and formant frequencies in distinguishing the voices of men and women. *Attention, Perception, & Psychophysics, 71*(5), 1150–1166. https://doi.org/10.3758/APP.71.5.1150

Hu, H., Dietz, M., Williges, B., & Ewert, S. D. (2018). Better-ear glimpsing with symmetrically-placed interferers in bilateral cochlear implant users. *The Journal of the Acoustical Society of America, 143*(4), 2128–2141. https://doi.org/10.1121/1.5030918

Jakien, K. M., Kampel, S. D., Gordon, S. Y., & Gallun, F. J. (2017). The benefits of increased sensation level and bandwidth for spatial release from masking. *Ear and Hearing, 38*(1), e13–e21. https://doi.org/10.1097/AUD.0000000000000352

Johnstone, P. M., & Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical Society of America, 120*(4), 2177–2189. https://doi.org/10.1121/1.2225416

Kidd, G., Jr., Mason, C. R., Best, V., & Marrone, N. (2010). Stimulus factors influencing spatial release from speech-on-speech masking. *The Journal of the Acoustical Society of America, 128*(4), 1965–1978. https://doi.org/10.1121/1.3478781

Kidd, G., Jr., Mason, C. R., Rohtla, T. L., & Deliwala, P. S. (1998). Release from masking due to the spatial separation of sources in the identification of nonspeech auditory patterns. *The Journal of the Acoustical Society of America, 104*(1), 422–431. http://doi.org/10.1121/1.423246

Kidd, G., Jr., Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., & Best, V. (2016). Determining the energetic and informational components of speech-on-speech masking. *The Journal of the Acoustical Society of America, 140*(1), 132–144. https://doi.org/10.1121/1.4954748

Liang, Z. A. (1963). The auditory perception of Mandarin tones. *Acta Physica Sinica, 26,* 85–91.

Marrone, N., Mason, C. R., & Kidd, G., Jr. (2008). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *The Journal of the Acoustical Society of America, 124*(5), 3064–3075. https://doi.org/10.1121/1.2980441

Moore, B. C. J., & Gockel, H. E. (2012). Properties of auditory stream formation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1591), 919–931. https://doi.org/10.1098/rstb.2011.0355

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. *The Journal of the Acoustical Society of America, 95*(2), 1085–1099. https://doi.org/10.1121/1.408469

Noble, W., & Perrett, S. (2002). Hearing speech against spatially separate competing speech versus competing noise. *Perception & Psychophysics, 64,* 1325–1336. https://doi.org/10.3758/bf03194775

Oh, Y., Hartling, C. L., Srinivasan, N. K., Eddolls, M., Diedesch, A. C., Gallun, F. J., & Reiss, L. A. J. (2019). Broad binaural fusion impairs segregation of speech based on voice pitch differences in a 'cocktail party' environment. bioRxiv 805309; https://doi.org/10.1101/805309

Poon, M. S. F., & Ng, M. L. (2015). The role of fundamental frequency and formants in voice gender identification. *Speech, Language and Hearing, 18*(3), 161–165. https://doi.org/10.1179/2050572814Y.0000000058

Rennies, J., Best, V., Roverud, E., & Kidd, G., Jr. (2019). Energetic and informational components of speech-on-speech masking in binaural speech intelligibility and perceived listening effort. *Trends in Hearing, 23,* 2331216519854597. https://doi.org/10.1177/2331216519854597

Rothauser, E. H., Chapman, W. D., Guttman, N., Nordby, K. S., Silbiger, H. R., Urbanek, G. E., & Weinstock, M. (1969). IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics, 17*(3), 227–246. https://doi.org/10.1109/IEEESTD.1969.7405210

Srinivasan, N. K., Jakien, K. M., & Gallun, F. J. (2016). Release from masking for small spatial separations: Effects of age and

hearing loss. *The Journal of the Acoustical Society of America, 140*(1), EL73–EL78. https://doi.org/10.1121/1.4954386

Srinivasan, N. K., Stansell, M., & Gallun, F. J. (2017). The role of early and late reflections on spatial release from masking: Effects of age and hearing loss. *The Journal of the Acoustical Society of America, 141*(3), EL185–EL191. https://doi.org/10.1121/1.4973837

Tao, D.-D., Fu, Q.-J., Galvin, J. J., III, & Yu, Y.-F. (2017). The development and validation of the Closed-Set Mandarin Sentence (CMS) test. *Speech Communication, 92,* 125–131. https://doi.org/10.1016/j.specom.2017.06.008

Tao, D.-D., Liu, Y.-W., Yei, F., Galvin, J. J., III, Chen, B., & Fu, Q.-J. (2018). Effects of age and duration of deafness on Mandarin speech understanding in competing speech by normal-hearing and cochlear implant children. *The Journal of the Acoustical Society of America, 144*(2), EL131–EL137. https://doi.org/10.1121/1.5051051

Vestergaard, M. D., Fyson, N. R. C., & Patterson, R. D. (2009). The interaction of vocal characteristics and audibility in the recognition of concurrent syllables. *The Journal of the Acoustical Society of America, 125*(2), 1114–1124. https://doi.org/10.1121/1.3050321

Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free-field listening: II. Psychophysical validation. *The Journal of the Acoustical Society of America, 85*(2), 868–878. https://doi.org/10.1121/1.397558

Wu, X., Wang, C., Chen, J., Qu, H., Li, W., Wu, Y., Schneider, B. A., & Li, L. (2005). The effect of perceived spatial separation on informational masking of Chinese speech. *Hearing Research, 199*(1–2), 1–10. https://doi.org/10.1016/j.heares.2004.03.010

Wu, X., Yang, Z., Huang, Y., Chen, J., Li, L., Daneman, M., & Schneider, B. A. (2011). Cross-language differences in informational masking of speech by speech: English versus Mandarin Chinese. *Journal of Speech, Language, and Hearing Research, 54*(6), 1506–1524. https://doi.org/10.1044/1092-4388(2011/10-0282)

Xie, X., & Myers, E. (2015). The impact of musical training and tone language experience on talker identification. *The Journal of the Acoustical Society of America, 137*(1), 419–432. https://doi.org/10.1121/1.4904699

Yuen, K. C. P., & Yuan, M. (2014). Development of spatial release from masking in Mandarin-speaking children with normal hearing. *Journal of Speech, Language, and Hearing Research, 57*(5), 2005–2023. https://doi.org/10.1044/2014_JSLHR-H-13-0060

Zobel, B. H., Wagner, A., Sanders, L. D., & Başkent, D. (2019). Spatial release from informational masking declines with age: Evidence from a detection task in a virtual separation paradigm. *The Journal of the Acoustical Society of America, 146*(1), 548–566. https://doi.org/10.1121/1.5118240