The Effect of Simulated Interaural Frequency Mismatch on Speech Understanding and Spatial Release From Masking

Matthew J. Goupell,[1] Corey A. Stoelb,[2] Alan Kan,[2] and Ruth Y. Litovsky[2]

Objective: The binaural-hearing system interaurally compares inputs, which underlies the ability to localize sound sources and to better understand speech in complex acoustic environments. Cochlear implants (CIs) are provided in both ears to increase binaural-hearing benefits; however, bilateral CI users continue to struggle with understanding speech in the presence of interfering sounds and do not achieve the same level of spatial release from masking (SRM) as normal-hearing listeners. One reason for diminished SRM in CI users could be that the electrode arrays are inserted at different depths in each ear, which would cause an interaural frequency mismatch. Because interaural frequency mismatch diminishes the salience of interaural differences for relatively simple stimuli, it may also diminish binaural benefits for spectral-temporally complex stimuli like speech. This study evaluated the effect of simulated frequency-to-place mismatch on speech understanding and SRM.

Design: Eleven normal-hearing listeners were tested on a speech understanding task. There was a female target talker who spoke five-word sentences from a closed set of words. There were two interfering male talkers who spoke unrelated sentences. Nonindividualized head-related transfer functions were used to simulate a virtual auditory space. The target was presented from the front (0°), and the interfering speech was either presented from the front (colocated) or from 90° to the right (spatially separated). Stimuli were then processed by an eight-channel vocoder with tonal carriers to simulate aspects of listening through a CI. Frequency-to-place mismatch ("shift") was introduced by increasing the center frequency of the synthesis filters compared with the corresponding analysis filters. Speech understanding was measured for different shifts (0, 3, 4.5, and 6 mm) and target-to-masker ratios (TMRs: +10 to –10 dB). SRM was calculated as the difference in the percentage of correct words for the colocated and separated conditions. Two types of shifts were tested: (1) bilateral shifts that had the same frequency-to-place mismatch in both ears, but no interaural frequency mismatch, and (2) unilateral shifts that produced an interaural frequency mismatch.

Results: For the bilateral shift conditions, speech understanding decreased with increasing shift and with decreasing TMR, for both colocated and separate conditions. There was, however, no interaction between shift and spatial configuration; in other words, SRM was not affected by shift. For the unilateral shift conditions, speech understanding decreased with increasing interaural mismatch and with decreasing TMR for both the colocated and spatially separated conditions. Critically, there was a significant interaction between the amount of shift and spatial configuration; in other words, SRM decreased for increasing interaural mismatch.

Conclusions: A frequency-to-place mismatch in one or both ears resulted in decreased speech understanding. SRM, however, was only affected in conditions with unilateral shifts and interaural frequency mismatch. Therefore, matching frequency information between the ears provides listeners with larger binaural-hearing benefits, for example, improved speech understanding in the presence of interfering talkers. A clinical procedure to reduce interaural frequency mismatch when programming bilateral CIs may improve benefits in speech segregation that are due to binaural-hearing abilities.

Key words: Binaural hearing, Cochlear implant, Sound localization, Spatial release from masking.

INTRODUCTION

Normal-hearing (NH) listeners use binaural cues to understand speech in complex auditory environments where there are multiple interfering sound sources (e.g., Bernstein et al. 2016; Brungart et al. 2001; Cherry 1953; Darwin & Hukin 2000; Freyman et al. 2001; Hawley et al., 2004). One way to assess the binaural hearing advantage for understanding speech in the presence of competing sounds is to measure spatial release from masking (SRM), or the difference in speech understanding when target and interferer are colocated and spatially separated. SRM is produced by several mechanisms (see Litovsky et al. 2006; Loizou et al. 2009 for an overview). One component results from "head shadow," which occurs when there is a better target-to-masker ratio (TMR) at one ear and does not necessitate binaural processing (i.e., it is monaural in nature). Another component results from binaural processing, which is sometimes called "squelch," and causes a binaural unmasking of speech. Squelch is the intelligibility benefit received when one has access to a second ear for listening, even though it has a poorer TMR compared with the contralateral ear. It is thought that the binaural benefit obtained via squelch results partially from the sounds being perceived at different locations in space. Unfortunately, people with hearing loss have diminished binaural unmasking abilities for a variety of reasons (e.g., Best et al. 2013; Misurelli & Litovsky 2015). Therefore, it is not surprising that a major complaint from people with hearing loss is communicating in noisy environments and complex auditory scenes (Hallberg et al. 2008; Kramer et al. 1998).

One of the long-term goals of intervention is to provide individuals with hearing loss with increased hearing functionality. Among those functional abilities are ones that require access to appropriately encoded binaural cues. Within the population with hearing loss, those with severe-to-profound hearing loss are candidates for cochlear implants (CIs). An increasing number of CI users undergo bilateral implantation (Peters et al. 2010) partially because of improved sound localization and speech understanding in background noise. Bilateral CI users demonstrate binaural-hearing benefits because they perform better at binaural-hearing tasks than unilateral CI users (Litovsky et al. 2006, 2012). Bilateral CI users demonstrate a range of SRM, from none to about 5 dB, possibly depending on the testing conditions and particular listeners (Bernstein et al. 2016; Buss et al. 2008;

[1]Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland, USA; and [2]Waisman Center, University of Wisconsin, Madison, Wisconsin, USA.

<zdoi; 10.1097/AUD.0000000000000541>

Goupell et al. 2016; Litovsky et al. 2006; Loizou et al. 2009; Mueller et al. 2002; van Hoesel & Tyler 2003; van Hoesel et al. 2008). For example, Loizou et al. (2009) found up to 6 dB of head shadow in bilateral CI listeners, but no binaural unmasking of speech in an SRM experiment where masker type and number of maskers were varied. Hawley et al. (2004), testing the same conditions in NH listeners, found the same 6 dB of head shadow, but much larger (8 dB) effect of binaural unmasking of speech. Currently, bilateral CI listeners still do not perform as well as their NH counterparts, even after the acoustic signals presented to the NH listeners are degraded with a CI simulation (i.e., vocoding) (Bernstein et al. 2016; Goupell et al. 2016).

The relatively small amount of binaural unmasking of speech in bilateral CI listeners likely results from a combination of device-related, surgical, and biological factors (for reviews, see Kan & Litovsky 2015; Litovsky et al. 2012). The purpose of this study is to focus on the effects of frequency-to-place mismatch and interaural frequency mismatch on binaural unmasking of speech, which can be caused by the type of array (a device-related factor), placement of the CI array in the cochlea (a surgical factor), or neural survival (a biological factor). While numerous factors could cause interaural frequency mismatch, binaural processing is most effective for frequency-matched inputs. Interaural frequency mismatch would reduce the salience of binaural cues and thus potentially limit the binaural unmasking of speech component of SRM. For this study, we will make a simplifying assumption and primarily consider interaural frequency mismatch caused by differences in insertion depth.

The human cochlea is on average approximately 35 mm in length, and electrode insertion depths can vary greatly from about 20 to 31 mm (Gstoettner et al. 1999; Helbig et al. 2012; Ketten et al. 1998). The relatively shallow insertion depth of an electrode array compared with the frequencies allocated to the electrodes typically introduces a frequency-to-place mismatch of information, meaning that CIs typically present relatively low-frequency information to neurons with relatively high characteristic frequencies (Landsberger et al. 2015). Relatively shallow insertion depths that cause frequency-to-place mismatch are thought to decrease speech understanding, at least initially, although some adaptation seems to occur (Skinner et al. 1995; Svirsky et al. 2004). The issue concerning bilateral implantation presents additional complications. Currently, there is no method for ensuring that electrode insertion depth is matched between the ears. Even if there was a method to match electrode insertion depths, the surgery is difficult and it would likely have small interaural differences (e.g., Pearl et al. 2014). Therefore, some amount of interaural frequency mismatch from differing insertion depths likely occurs for bilateral CI users. Psychophysical studies on interaural frequency mismatch suggest that implantation depths between the electrode arrays may vary by several millimeters (Kan et al. 2013, 2015; Laback et al. 2004; Long et al. 2003; Poon et al. 2009; van Hoesel & Clark 1997).

The effects of interaural frequency mismatch have been investigated in several experimental paradigms in bilateral CI and NH listeners. Seminal studies on the topic used amplitude-modulated high-frequency tones or narrow bands of noise presented to NH listeners (Blanks et al. 2007, 2008; Francart & Wouters 2007; Henning 1974; Nuetzel & Hafter 1981); as the overlap in the frequency spectrum became smaller, interaural time difference (ITD) or interaural level difference (ILD)

discrimination sensitivity was reduced. In a more direct comparison of bilateral CI to NH listeners, binaural processing from single-electrode monopolar stimulation from pitch-matched pairs (Goupell 2015; Kan et al. 2013, 2015) was compared with bandlimited pulse trains designed to mimic aspects of single-electrode monopolar stimulation (Goupell 2015; Goupell et al. 2013). The conclusion from those comparisons was that 3 mm of interaural frequency mismatch significantly decreased ITD and ILD discrimination sensitivity, reduced binaural fusion of stimuli, and distorted and diminished intracranial ITD and ILD lateralization; the broad tuning that comes with the use of monopolar electrical stimulation (bandwidths up to about 5 mm; Nelson et al. 2008, 2011) provided some tolerance for interaural place-of-stimulation mismatch.

The effects of interaural frequency mismatch on speech understanding have also been tested using channel vocoded speech. Yoon et al. (2011) tested speech understanding using vocoded stimuli with varying degrees of interaural frequency mismatch. Testing was conducted in quiet or noise, however, without spatial differences and without explicit training. For interaural mismatches of ≤1 mm in magnitude, they found that speech understanding improved by about 10% when listening to two ears compared with one, which is consistent with the typical magnitude of binaural redundancy (Bronkhorst & Plomp 1988). They also found poorer speech understanding when there was a ≥2 mm simulated basalward shift. In a similar study, van Besouw et al. (2013) tested diotic speech perception in noise with simulated combinations of insertion depths, and each ear was varied independently from 17 to 26 mm simulated insertion depths. There was only 15 min of training, and no spatial differences were applied. They found better speech understanding when listening to two ears compared with one for conditions where insertion depths were deep (i.e., there was less frequency-to-place mismatch) and there was less interaural mismatch. For shallower-simulated insertion depths or larger interaural mismatches, speech understanding did not improve for two ears compared with one.

Siciliano et al. (2010) tested vocoded speech understanding in quiet using six channels that were interleaved, such that odd-numbered channels were presented to one ear and even-numbered channels were presented to the other. They then shifted the output frequency range of one ear basally by 6 mm. This study was novel in that they provided 10 hr of training, which is critical for decent speech understanding when listening to relatively large shifts (Rosen et al. 1999), and the interleaved processing manipulation provided a better possibility that the listeners could use the speech information in the shifted channels to their benefit, in contrast to the stimuli in Yoon et al. (2011) and van Besouw et al. (2013). Performance, however, never exceeded that of the unshifted ear alone, even though the shifted ear contained potentially useful speech information not presented to the unshifted ear. Therefore, it seems that there is an upper bound on how much the speech information can be shifted before it becomes unusable, particularly in the presence of unshifted, more intelligible speech information.

Another method of assessing the benefit of having an ear with frequency-to-place mismatch for understanding speech is to place target and maskers at different spatial locations, conditions that were omitted from the previously mentioned vocoder studies. In other words, investigating the effects of frequency-to-place mismatch or interaural mismatch on SRM may better

show how the information across the two ears is processed together, which was the purpose of this study. Yoon et al. (2013) studied the effects of frequency-to-place mismatch or interaural mismatch on SRM using sentences as the target and noise as the masker at +5- and +10-dB TMR. Unlike Rosen et al. (1999) and Siciliano et al. (2010), no training was provided. They found that binaural unmasking diminished with mismatch but head shadow did not. There are, however, several reasons to reconsider looking at the effects of frequency-to-place mismatch or interaural mismatch on SRM. First, SRM and binaural unmasking in particular are largest for negative TMRs (Bernstein et al. 2016; Freyman et al. 2008). Second, the amount of SRM and binaural unmasking depends critically on the target and masker similarity; the largest binaural benefits are observed with speech maskers on speech targets, conditions that demonstrate a large amount of informational masking (Bernstein et al. 2016; Brungart 2001; Durlach et al. 2003). Third, training is critical for understanding vocoded speech with relatively large amounts of frequency-to-place mismatch (Rosen et al. 1999; Waked et al. 2017). Therefore, we performed the present study addressing these three design choices to maximize the SRM and potentially the effect of mismatch.

We hypothesized that SRM decreases with increasing interaural mismatch (i.e., unilateral shifts). This is because interaural mismatch decreases binaural sensitivity, reduces binaural fusion, and distorts and diminishes perceived lateralization for ITDs and ILDs (Goupell 2015; Goupell et al. 2013; Kan et al. 2013, 2015). One potential confound in performing such a study, however, is whether the effects can be solely attributed to an interaural frequency mismatch or result from poorer unilateral speech understanding from frequency-to-place mismatch (Dorman et al. 1997; Rosen et al. 1999; Shannon et al. 1998). Therefore, bilateral shift conditions with no interaural frequency mismatch were also tested. For the bilateral shift conditions, we hypothesized that SRM would be independent of frequency-to-place mismatch for interaurally matched carriers because the binaural information that produces SRM would be conveyed to frequency-matched inputs.

MATERIALS AND METHODS

Listeners

Eleven listeners (17 to 29 years) participated in the experiment. Listeners had thresholds that were less than 20 dB HL measured at octave frequencies from 250 to 8000 Hz, and thresholds did not differ by more than 15 dB between ears. Listeners signed a consent form before participation and were paid an hourly wage for their time. Some listeners had previous experience in psychoacoustic experiments. None of the listeners had any previous experience listening to vocoded speech.

Stimuli and Equipment

Stimuli for the experiment were generated using MATLAB (MathWorks, Natick, MA) and delivered via a Tucker-Davis Technologies System 3 (RP2.1, PA5, HB7; Alachua, FL) to insert-ear headphones (ER2; Etymotic, Elk Grove Village, IL). Testing was completed in a sound-attenuating booth.

The targets consisted of five-word sentences spoken by a female. Each word was chosen at random from a separate set of eight words to form a syntactically coherent sentence, although sentences were not necessarily semantically coherent

(Kidd et al. 2008). The structure of the sentence was a matrix sentence test (name, verb, number, adjective, plural object). For example, "Jill bought three blue pens." The masker consisted of two spatially colocated male talkers speaking streams of randomly selected low-context sentences (IEEE; Rothauser et al. 1969). All speech was filtered through a set of nonindividualized head-related transfer functions (HRTFs) from the CIPIC database (KEMAR with large pinna) to provide azimuthal spatial cues (Algazi et al. 2001). Target speech was filtered with front location (0°) HRTFs. The masker was filtered with a front location (0°) HRTF for the colocated condition or right side (+90°) HRTF for the spatially separated condition. The target and maskers were mixed at four different TMRs = +10, 0, −5, and −10 dB. The TMR was manipulated by holding the level of the masker constant at 60 dBA and adjusting the level of the target speech.

To simulate some of the aspects of CI sound processing, including the limited frequency resolution and the omission of temporal fine structure, an eight-channel vocoder was used to process signals that were delivered to the NH listeners. The use of vocoders to simulate some of the aspects of CI processing is a commonly used tool in the field of CI research (Loizou 2006; Shannon et al. 1995); however, it is well known that vocoders are not perfect simulations of how CI functions (e.g., Aguiar et al. 2016). Vocoder simulations provide an opportunity to examine variability related to the factor of interest (i.e., insertion depth and electrode array placement) in the absence of factors that are known to vary with individual subjects (e.g., device parameter setting, neural integrity, language development). Vocoder simulations provide the flexibility that one does not have to rely on the subject-specific array locations and neural health; it is not possible to experimentally move the arrays physically in the cochlea. Such subject variability in the CI population, compounded by the issue of using spectral-temporally complex speech signals, makes testing of frequency-to-place mismatch and interaural mismatch practically untenable.

Vocoder processing was performed after HRTF filtering and mixing of target and maskers. Eight contiguous fourth-order band-pass Butterworth filters were used and had corner frequencies that were logarithmically spaced from 150 to 8000 Hz. Envelope information was then extracted via a second-order low-pass Butterworth filter with a 50-Hz cutoff frequency. A 50-Hz cutoff was selected to limit resolved sidebands for some of the channels and conditions, which generally improve word recognition scores (Souza & Rosen 2009) and would have confounded interpretation of the data. Each envelope was then used to modulate a sinusoidal carrier with a frequency equal to the center frequency of its respective analysis band or systematically shifted to higher frequencies in conditions that required a spectral shift. The final step in the vocoding process was to sum the envelope-modulated sine waves.

A sinusoidal carrier was used instead of a narrowband noise for two reasons. First, interaural frequency shifting with narrowband noises and concomitantly changing the frequency content across the ears decrease the interaural cross-correlation of each band (Bernstein & Trahiotis 1996; Goupell 2010; Goupell et al. 2013). NH listeners are very sensitive to changes in the interaural correlation (Gabriel & Colburn 1981; Goupell & Litovsky 2014) because it creates a diffuse spatial percept of the sound image (Blauert & Lindemann 1986; Whitmer et al. 2012). Avoiding narrowband noise carriers by using sinusoidal carriers

minimizes the diffuse spatial percept generated by the carriers. Second, the vocoder is intended to convey information about the temporal envelope to the listener. Narrowband noises, however, introduce envelope fluctuations (Whitmal et al. 2007) that are guaranteed to be different across the ears for the mismatched conditions; those differences increase for greater amounts of interaural mismatch, thus corrupting the main acoustic information that we are attempting to convey to the listener.

The effect of interaurally matched and mismatched frequency shifts on speech understanding and SRM was examined. Upward frequency shifting was introduced by increasing the frequency of the sine wave carrier in each channel. The spectral shifts were transformed into millimeter units to reflect anatomical shift in the place of stimulation (Greenwood 1990). In experiment 1, both ears were shifted. This experiment was performed because one potential confound was whether the effects could be attributed to an interaural frequency mismatch or poorer speech understanding from frequency-to-place mismatch (Dorman et al. 1997; Rosen et al. 1999; Shannon et al. 1998). Bilateral frequency-to-place mismatches of 0, 3, 4.5, and 6 mm were tested.

In experiment 2, only one ear was shifted to a higher frequency range leading to an interaural frequency mismatch (Δ), and the simulated mismatches were $\Delta = 0$, ± 3, ± 4.5, and ± 6 mm (note the $\pm$ symbols, as the carriers in either the left or right ear were increased in frequency). Positive values of mismatch indicate an increase in carrier frequency of the right ear, and negative values of mismatch indicate an increase in carrier frequency of the left ear. Figure 1 shows the frequencies of each sine wave carrier for the stimuli. Note that the 3- and 6-mm shifts align seven and six carrier center frequencies, respectively; in other words, they represent one- and two-channel shifts.

Procedure

All testing was completed via a listener-controlled computer program. Listeners used the computer mouse or touch screen to select options displayed on a computer monitor. Listeners initiated each trial with a button press. After hearing the stimulus, they selected the five words from an 8×5 matrix of words that included the eight possible options for each of the five-word categories. The order of testing was randomized across listeners. There were 32 conditions that
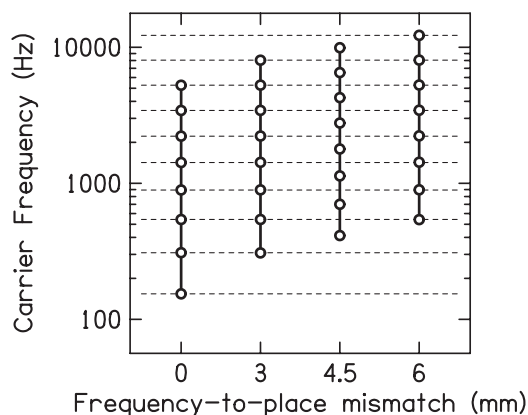

Fig. 1. Carrier frequencies of the eight vocoder channels for different amounts of frequency-to-place mismatch.

had bilateral shifts (4 TMRs × 4 Matched Shift Conditions × 2 Masker Locations) and 42 conditions that had unilateral shifts (3 TMRs × 7 Mismatched Shift Conditions × 2 Masker Locations) (only three TMRs were tested in experiment 2, the unilateral shift conditions, because of the near-floor effects observed at −10-dB TMR for the bilateral shift conditions.) For each condition, 15 sentences were tested. Therefore, a total of 1110 sentences were tested. Testing order was pseudorandomized and was divided into blocks. Testing time was approximately 10 hr and was divided into five 2-hr sessions.

Before the testing phase, there was a 2-hr training session to improve listeners' ability to understand frequency-mismatched vocoded speech. The task was the same for the training and testing phases. During training, the correct answer was shown on the screen and the entire stimulus was played again (Fu & Galvin 2003). During this feedback portion, the instructions were to listen for the correct answer that was highlighted on the screen. Training only occurred for colocated targets and maskers, the relatively more difficult of the conditions. The TMR vas varied pseudorandomly between +10 and 0 dB. Training was not performed at lower TMRs because vocoded speech understanding for colocated talkers at negative TMRs was too difficult for training purposes. Listeners were initially trained with no shift ($\Delta = 0$ mm) until they were comfortable listening to vocoded speech, then all the frequency shifts were trained by randomizing all the shifts in blocks, like what occurred in the testing phase. For all listeners, performance saturated within the 2-hr session, consistent with previous reports on training shifted vocoded speech with this specific speech corpus (Waked et al. 2017). Listeners were informally asked if the vocoded stimuli sounded as if they originated from a point in space (externalized), as opposed to a point inside the head (internalized). Listeners generally perceived the stimuli as externalized, but noted that higher degrees of mismatch seemed to disrupt the amount of externalization.

Data Analysis

The percentage of correct words was calculated for each condition. SRM was calculated for each condition as the difference between speech understanding scores when the maskers were spatially separated and when they were colocated with the target. Positive SRM values indicate better performance in the separated condition and a benefit for spatial separation between target and maskers. To minimize the inhomogeneity of variance that can occur with proportional data, a rationalized arcsine unit (RAU) transformation was performed on scores before statistical analysis (Studebaker 1985). Inferential statistics were performed using repeated-measures analyses of variance (RM ANOVAs). Post hoc analyses were performed using planned two-sample, two-tail, paired t-tests.

RESULTS

Effect of Frequency-to-Place Mismatch With Bilateral Shifts

The speech understanding scores in RAUs for the bilateral shift conditions are shown in Figures 2A–D, and amount of SRM is shown in Figures 2E–H. A three-way RM ANOVA with
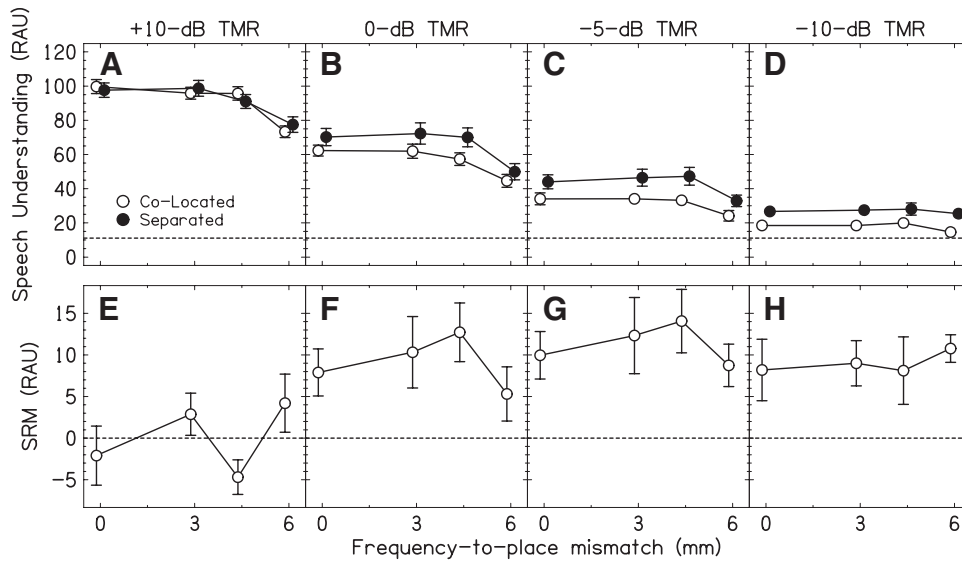
Fig. 2. Speech understanding and SRM for conditions with bilateral shifts and no interaural frequency mismatch. A–D, Average speech understanding scores in RAUs for colocated (open symbols) and separated (closed symbols) conditions. The horizontal dashed line represents chance performance. E–H, Average SRM for the bilateral shift conditions. Error bars in all panels are ±1 SE in length. RAU indicates rationalized arcsine unit; SRM, spatial release from masking; TMR, target-to-masker ratio.

factors location, Δ, and TMR was performed. Speech understanding decreased with decreasing TMR [$F(3,30) = 261.4$, $p < 0.001$, $\eta_p^2 = 0.96$]; post hoc tests showed that performance for all TMRs was different ($p < 0.001$ for all). Speech understanding increased from the colocated to separated conditions [$F(1,10) = 15.7$, $p = 0.003$, $\eta_p^2 = 0.61$]. Speech understanding decreased with increasing Δ [$F(3,30) = 57.9$, $p < 0.001$, $\eta_p^2 = 0.85$]. There was a significant TMR × location interaction [$F(3,30) = 11.3$, $p < 0.001$, $\eta_p^2 = 0.53$]. There was a significant TMR × Δ interaction [$F(9,90) = 8.84$, $p < 0.001$, $\eta_p^2 = 0.47$]. Critically, there was no significant location × Δ interaction ($p > 0.05$), which is best observed as the relatively flat lines in Figures 2E–H. This outcome is consistent with our hypothesis that frequency-to-place mismatch from a bilateral shift that does not introduce interaural frequency mismatch should not affect the amount of SRM. The location × TMR × Δ interaction was also not significant ($p > 0.05$).

To better investigate the significant TMR × location and TMR × Δ interactions in the three-way RM ANOVA, we performed four separate two-way RM ANOVAs with factors location and Δ, one for each TMR. For +10-dB TMR, the effect of location was not significant [$F(1,10) = 0.001$, $p = 0.91$, $\eta_p^2 = 0.00$], the effect of Δ was significant [$F(3,30) = 33.8$, $p < 0.001$, $\eta_p^2 = 0.77$], and the location × Δ interaction was not significant [$F(3,30) = 2.37$, $p = 0.091$, $\eta_p^2 = 0.19$]. This means that there was no improvement in speech understanding when spatial separation was introduced (i.e., no SRM, see Fig. 2E) and speech understanding decreased with increasing Δ. For 0-dB TMR, the effect of location was significant [$F(1,10) = 10.4$, $p = 0.009$, $\eta_p^2 = 0.51$], the effect of Δ was significant [$F(3,30) = 32.2$, $p < 0.001$, $\eta_p^2 = 0.76$], and the location × Δ interaction was not significant [$F(3,30) = 1.69$, $p = 0.19$, $\eta_p^2 = 0.15$]. This means that there was an improvement in speech understanding when spatial separation was introduced (see Fig. 2F) and speech understanding decreased with increasing Δ. For −5-dB TMR, the effect of location was significant [$F(1,10) = 24.8$, $p = 0.001$, $\eta_p^2 = 0.71$], the effect of

Δ was significant [$F(3,30) = 16.5$, $p < 0.001$, $\eta_p^2 = 0.62$], and the location × Δ interaction was not significant [$F(3,30) = 0.58$, $p = 0.64$, $\eta_p^2 = 0.05$]. This means that, like 0-dB TMR, there was an improvement in speech understanding when spatial separation was introduced (see Fig. 2G) and speech understanding decreased with increasing Δ. For −10-dB TMR, the effect of location was significant [$F(1,10) = 20.3$, $p = 0.001$, $\eta_p^2 = 0.67$], the effect of Δ was not significant [$F(3,30) = 2.36$, $p = 0.092$, $\eta_p^2 = 0.19$], and the Location × Δ interaction was not significant [$F(3,30) = 0.19$, $p = 0.90$, $\eta_p^2 = 0.02$]. This means that there was an improvement in speech understanding when spatial separation was introduced (see Fig. 2H), like the 0- and −5-dB TMRs, and speech understanding did not change with increasing Δ, unlike the other TMRs. In summary, (1) SRM did not occur for the +10-dB TMR but did for the lower TMRs (best seen in Figs. 2E–H), and (2) speech understanding decreased with increasing Δ except at −10-dB TMR, which may be a result of a floor effect (best seen in Figs. 2A–D).

Effect of Interaural Frequency Mismatch With Unilateral Shifts

The speech understanding scores in RAUs for the unilateral shift conditions are shown in Figures 3A–C, and the amount of SRM is shown in Figures 3D–F. Remember that the separated conditions move the interferers +90° to the right of the listener and that $\Delta > 0$ introduce shift in the right ear. A three-way RM ANOVA with factors Δ, location, and TMR was performed. Speech understanding decreased with decreasing TMR [$F(2,20) = 371.0$, $p < 0.001$, $\eta_p^2 = 0.97$]; post hoc tests showed that performance for all TMRs was different ($p < 0.001$ for all). Speech understanding increased from the colocated to separated conditions [$F(1,10) = 10.3$, $p = 0.009$, $\eta_p^2 = 0.51$]. Speech understanding decreased with increasing magnitude of Δ [$F(6,60) = 43.5$, $p < 0.001$, $\eta_p^2 = 0.81$]. There was a significant TMR × Δ interaction [$F(12,120) = 2.40$, $p = 0.008$, $\eta_p^2 = 0.19$]. Most importantly, there was a significant location × Δ interaction
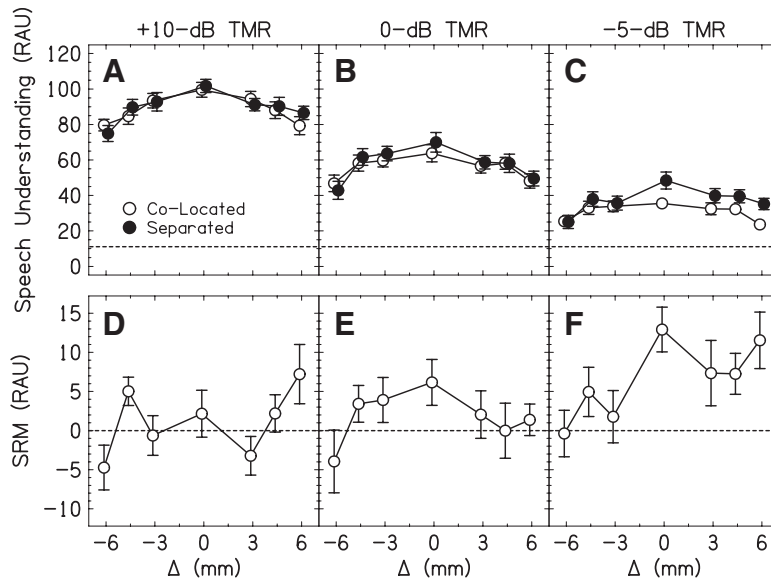
Fig. 3. Speech understanding and SRM for conditions with unilateral shifts and interaural frequency mismatch. A–C, Average speech understanding scores in RAUs for colocated (open symbols) and separated (closed symbols) conditions. The horizontal dashed line represents chance performance. D–F, Average SRM for the unilateral shift conditions. Error bars in all panels are ±1 SE in length. RAU indicates rationalized arcsine unit; SRM, spatial release from masking; TMR, target-to-masker ratio.

such that the largest increase in RAU from colocated to separated locations mostly decreased with increasing Δ [$F(6,60) = 5.01, p = 0.001, \eta_p^2 = 0.33$], which is broadly consistent with our main hypothesis. The location × TMR and location × TMR × Δ interactions were not significant ($p > 0.05$).

To better investigate the significant Δ × TMR interaction in the three-way RM ANOVA, we performed three separate two-way RM ANOVAs with factors location and Δ, one for each TMR. For the +10-dB TMR, the effect of location was not significant [$F(1,10) = 0.04, p = 0.54, \eta_p^2 = 0.04$], the effect of Δ was significant [$F(6,60) = 25.1, p < 0.001, \eta_p^2 = 0.72$], and the location × Δ interaction was significant [$F(6,60) = 3.63, p = 0.004, \eta_p^2 = 0.27$]. Subsequent post hoc tests showed that the separated condition was higher than the colocated condition for the −4.5-mm Δ ($p = 0.010$). For the 0-dB TMR, the effect of location was not significant [$F(1,10) = 1.09, p = 0.32, \eta_p^2 = 0.10$], the effect of Δ was significant [$F(6,60) = 26.1, p < 0.001, \eta_p^2 = 0.72$], and the location × Δ interaction was not significant [$F(6,60) = 1.49, p = 0.20, \eta_p^2 = 0.13$]. This means that speech understanding decreased with increasing Δ. For −5-dB TMR, the effect of location was significant [$F(1,10) = 10.2, p = 0.010, \eta_p^2 = 0.51$], the effect of Δ was significant [$F(6,60) = 15.4, p < 0.001, \eta_p^2 = 0.61$], and the location × Δ interaction was significant [$F(6,60) = 3.01, p = 0.012, \eta_p^2 = 0.23$]. Subsequent post hoc tests showed that the separated condition was higher than the colocated condition for the 0-mm Δ ($p = 0.001$), +4.5-mm Δ ($p = 0.020$), and +6-mm Δ ($p = 0.009$).

In summary, the amount of SRM was a nonmonotonic function of Δ for the unilateral shift conditions (Figs. 3D–F). Contrary to expectations, SRM did not always decrease systematically for increasing mismatch.

Comparison of Bilateral and Unilateral Shifts

To better understand the nonmonotonicities in the data in Figures 2 and 3, we performed a comparison of the bilateral and unilateral shift conditions. Figure 4 plots the difference in speech understanding in RAUs for the bilateral (Fig. 2) and unilateral shift conditions (Fig. 3). Remember that ±Δs were tested for the unilateral shift conditions, but only $\Delta \geq 0$ were tested for the bilateral shift conditions. To accommodate the unequal number of conditions, the difference between the $\Delta < 0$ unilateral shifts were taken from the positive bilateral shift (e.g., the speech understanding score difference was calculated between the $\Delta = -6$-mm unilateral shift and $\Delta = +6$-mm bilateral shift conditions). If this speech understanding score difference was positive, then the bilateral shift condition where both ears have frequency-to-place mismatch and better access to interaurally matched inputs provided better speech understanding than the condition with only one ear with frequency-to-place mismatch.

Figure 4A shows the difference between the bilateral and unilateral shift conditions as a function of Δ for the colocated talkers at +10-dB TMR. As expected, when $\Delta = 0$ mm, the difference in speech understanding was zero because there was no difference in the stimuli. The overall function was symmetrical about $\Delta = 0$ mm. As the magnitude of Δ increased, the bilateral shift conditions provided better speech understanding, up to about 10 RAU for $\Delta = \pm 4.5$ mm. For $\Delta = \pm 3$ mm, which is a one-channel shift (see Fig. 1), there was an emerging benefit for the bilateral shift condition, perhaps because of the incongruence between the two ears for the unilateral shift conditions. For $\Delta = \pm 4.5$ mm, which is a 1.5-channel shift (see Fig. 1), this possible incongruence for the unilateral shift conditions was even greater. It could also be the case that the carriers were interleaved for this condition and this contributed to the relatively large scores for the bilateral shift conditions. Finally, for $\Delta = \pm 6$ mm, which is a two-channel shift (see Fig. 1), the bilateral shift conditions provided worse speech understanding than the unilateral shift conditions, the opposite direction of the smaller in magnitude Δs. This suggests that the listener may have attended to only the unshifted ear in the unilateral condition, similar to the results of Siciliano et al. (2010).
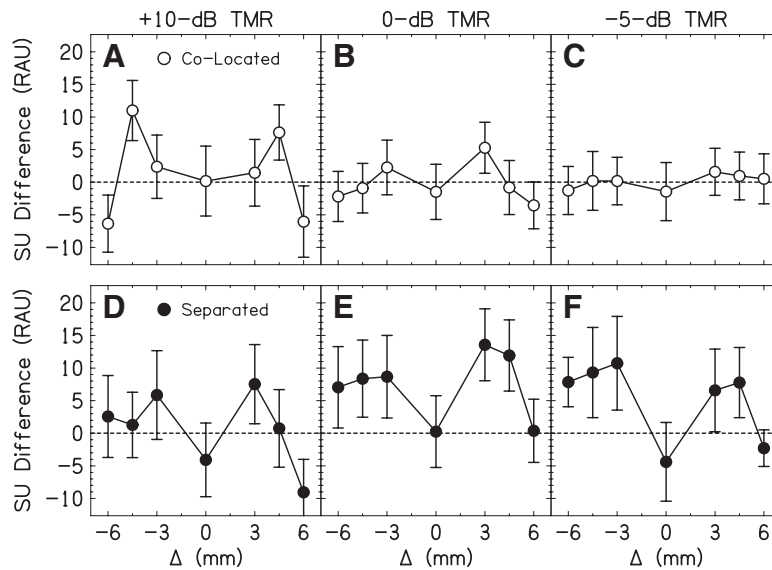
Fig. 4. Difference in speech understanding in RAUs between the bilateral and unilateral shift conditions. Error bars are ±1 SE in length. RAU indicates rationalized arcsine unit; SU, speech understanding; TMR, target-to-masker ratio.

Figures 4B, C show the colocated conditions for the 0- and −5-dB TMRs, respectively. They show the same symmetric functions as in Figure 4A; however, the difference between the bilateral and unilateral shift conditions was diminished. By −5-dB TMR (Fig. 4C), there was essentially no difference between the bilateral and unilateral shift conditions.

Figure 4D shows the difference between the bilateral and unilateral shift conditions as a function of Δ for the spatially separated talkers at +10-dB TMR. Like Figure 4A, when $\Delta = 0$ mm, the difference in speech understanding was zero because there was no difference in the stimuli. Unlike Figure 4A, the function peaked at $\Delta = \pm 3$ mm, not ± 4.5 mm. Also, the function was not symmetrical about $\Delta = 0$ mm, specifically at $\Delta = \pm 6$ mm. The changing location of the peak of the function might be explained by the incongruent carriers for the ± 4.5-mm unilateral shift condition (see Fig. 1). Because the carriers were mismatched, the binaural processing was relatively poor and little SRM could be achieved. The $\Delta = +6$-mm condition shows the same drop in the bilateral versus unilateral shift difference, similar to what occurred in colocated condition for Figure 4A. This drop did not occur at $\Delta = -6$-mm condition. This may have occurred because there was little SRM at +10-dB TMR and the ear with the poorer TMR was the unshifted ear. Thus, it may have been easier for the listeners to attend to the shifted ear with the better TMR even if the speech understanding was relatively poor.

Figures 4E, F show the difference between the bilateral and unilateral shift conditions as a function of Δ for the spatially separated talkers at 0- and −5-dB TMRs, respectively. These functions were basically the same as in Figure 4D; however, all the points for $\Delta \neq 0$ mm were shifted upward, suggesting a larger benefit from spatial separation for the bilateral shift conditions. Interaural frequency mismatch diminished this benefit for increasing Δ when the right ear, the ear nearest the interferers, was shifted. At the $\Delta = +6$-mm condition, there was no difference between the bilateral and unilateral shift conditions. This might have occurred because the binaural benefit for the bilateral shift equaled the advantage for having much better speech understanding in the unshifted left ear. For $\Delta < 0$ mm,

the bilateral shift conditions may have provided better speech understanding because the binaural benefits for the matched carriers outweighed the poor speech understanding in the unilateral unshifted right ear, which had the poorer TMR.

In summary, the comparison between the bilateral and unilateral shift conditions help in understanding the nonmonotonic functions seen in Figures 2 and 3. For the colocated conditions, speech understanding was better for bilateral shifts with small Δs, when the speech was relatively intelligible and the speech information was fairly congruent across the two ears. Speech understanding was better for unilateral shifts and $\Delta = +6$ mm, when the listener seemed to attend to the unshifted ear alone. These benefits/strategies are largest for +10-dB TMR and diminished for decreasing TMRs. For the separated conditions, the speech understanding for the bilateral shifts was best for the 0- and −5-dB TMRs because this was where SRM occurred. For $\Delta = +6$ mm, spatial advantage disappeared, perhaps because of offsetting binaural benefits for the bilateral shift condition and the advantage for having an unshifted left ear for the unilateral shift condition. For $\Delta = -6$ mm, there was a spatial advantage for the bilateral shift, which may have been because the listener was attending to an unshifted right ear, but that ear provided little help because of the poor TMR.

DISCUSSION

There currently exists a performance gap in binaural processing abilities (sound localization and speech understanding in the presence of competing sounds) between bilateral CI listeners using their speech processors and NH listeners. Most studies show that the benefits in bilateral CI listeners are derived from monaural head shadow, with limited amounts of binaural unmasking or squelch (e.g., Loizou et al. 2009; Schleich et al. 2004; van Hoesel et al. 2008). One factor that could affect performance is the placement of the electrode arrays, which relates to frequency-to-place mismatch and interaural frequency mismatch. It has been well established that frequency-to-place mismatch diminishes speech understanding in vocoder simulations (Başkent & Shannon 2007; Dorman et al. 1997; Rosen et al.

1999; Shannon et al. 1998) and acute studies with CIs (Fu & Shannon 1999). The improvement in speech understanding over the first year after activation is attributed, at least in part, to plasticity and adaptation to frequency-to-place mismatch (Blamey et al. 2013; Fu & Shannon 1999), although there is some evidence to the contrary (Fu et al. 2002).

In patients with bilateral CIs, a different sort of mismatch occurs whereby the same cochlear place of stimulation in the two ears can receive mismatched frequency inputs. Interaural mismatches of this type decrease binaural sensitivity for simple stimuli (Goupell 2015; Goupell et al. 2013; Kan et al. 2013, 2015). A possible consequence of interaural frequency mismatch could therefore be diminished binaural processing and SRM for speech stimuli.

This study was designed to test the hypothesis that while frequency-to-place mismatch would have no effect on SRM, interaural frequency mismatch would cause decreasing SRM. To test this hypothesis in a situation that avoids individual subject factors such as neural health and unknown array placement, we simulated frequency-to-place mismatch and interaural frequency mismatch as would occur with CI processing and presented these stimuli to NH listeners.

Results showed that speech understanding decreased with increasing frequency-to-place mismatch for bilateral shifts (with no interaural mismatch, Fig. 2) and unilateral shifts (with interaural mismatch, Fig. 3), consistent with previous reports (Başkent & Shannon 2007; Dorman et al. 1997; Rosen et al. 1999; Shannon et al. 1998). SRM was calculated by comparing speech understanding for colocated talkers and spatially separated talkers, where the target was located in front of the listener and the interferer was located at 90° to the right of the listener. There was increased SRM as the TMR decreased, but bilateral shifts were not significantly affected by frequency-to-place mismatch (Figs. 2D–H). Relatively low TMRs demonstrated the largest SRM, which commonly occurs in SRM experiments because nonperfect performance is necessary for the listener to benefit from the perceived spatial separation (Bernstein et al. 2016; Brungart et al. 2001).

Unilateral shifts, on the other hand, significantly affected SRM through a significant $\Delta \times$ location interaction (Figs. 3D–F). Specifically, interaural frequency mismatches of $\Delta = 0$ (i.e., no mismatch) and $+6\,mm$ (i.e., the right electrode array shifted upward; see Fig. 1 for frequencies) had the largest SRM (Figs. 3D–F) and were significantly larger than SRM for $\Delta = -6\,mm$. This pattern differed from our hypothesis, where we expected decreasing SRM for increasing magnitude of Δ. Such a hypothesis is derived solely from the binaural unmasking or squelch effect. By $6\,mm$ of interaural frequency mismatch, the binaural information in the speech envelopes may be interaurally decorrelated enough to remove most differences in perceived spatial location. Nonetheless, the amount of SRM as a function of Δ observed in Figures 3D–F might be explained by observing the differences in speech understanding between the unilateral and bilateral shift conditions (Fig. 4). Head shadow should be minimal at low frequencies and increasingly large at higher frequencies (Feddersen et al. 1957). Therefore, the high-frequency channels in the left ear have access to the best TMRs. Leaving the left ear unshifted (i.e., no frequency-to-place mismatch) keeps the better ear more intelligible for $\Delta > 0$ (see Figs. 4D–F). The right ear becomes less intelligible with shift, which has relatively more interferer energy. By $\Delta = +6\,mm$, the interaural frequency mismatch was two channels (see Fig. 1),

much larger than the amount of mismatch needed to see decrements in binaural performance in other studies (Goupell 2015; Goupell et al. 2013; Kan et al. 2013, 2015). Note that the amount of SRM for $\Delta = +6\,mm$ at -10-dB TMR for the unilateral and bilateral shifts was approximately 11 RAUs. The difference between the $\Delta = +6\,mm$ and $-6\,mm$ conditions may be an indication that listeners attended to the unshifted left ear and tried to completely ignore the information in the shifted right ear. Such an interpretation would be similar to what was observed by Siciliano et al. (2010), where listeners ignored additional speech information in a shifted ear in the presence of speech information in an unshifted ear.

In the other direction, $\Delta = -6\,mm$ shifted frequency information in the left ear. Remember that the left ear is the better ear from head shadow because the interferers are placed at 90° to the right. The left ear, however, has decreased speech understanding from frequency-to-place mismatch. Therefore, listeners could try to glean unshifted speech information from the right ear, which has the poorer TMR. What is interesting is that even with better TMRs in the left ear, the frequency-to-place mismatch appeared to make the speech too difficult to understand well.

Our study can be compared with another study by Yoon et al. (2013) who also investigated SRM by bilateral and unilateral shifts in NH listeners presented vocoded speech. Yoon et al. found significant head shadow for matched carrier conditions and when the ear with less frequency-to-place mismatch had the better TMR, but the head shadow was greatly reduced when the ear with more frequency-to-place mismatch had the better TMR. They also found increased speech understanding (i.e., binaural redundancy) for matched conditions and decreased speech understanding for mismatched conditions. Only a small amount of binaural unmasking was observed for the matched conditions, likely because they used a noise masker instead of another talker (Bernstein et al. 2016). Large and positive binaural unmasking was observed for the mismatched conditions when the ear with less frequency-to-place mismatch was added; interference (i.e., a reduction in speech understanding from adding a second ear) was observed when the ear with more frequency-to-place mismatch was added. Binaural speech interference has also been observed in some bilateral CI listeners, which could be a result of a combination of factors, including differences in hearing history or insertion depth (Bernstein et al. 2016; Goupell et al. 2017).

Although we cannot separate the individual components of SRM, our data are broadly consistent with Yoon et al. (2013). We see less SRM and even negative SRM when the ear with the better TMR has the most frequency-to-place mismatch, the $\Delta = -6$-mm condition (Fig. 3). There are some noteworthy methodological differences between the two studies. One of the goals of this study was to maximize the binaural unmasking effect and the possibility that the listeners utilized information in the shifted ear. This was achieved in the present study by testing at negative TMRs and using a speech, rather than a noise, masker. Specifically, Yoon et al. used IEEE sentences as target speech and the interferer was a steady state speech-spectrum shaped noise. In contrast, we used a closed set of target words with IEEE sentences as interferers. They tested TMRs of $+5$ and $+10$ dB; in contrast, we tested TMRs of $+10$, 0, -5, and -10 dB to maximize informational masking and to achieve the largest SRM (Bernstein et al. 2016; Brungart et al. 2001). We also included more substantial training because training for shifted vocoded is critical for decent speech understanding (Rosen et al. 1999). Yoon et al. performed 30 min of

familiarization for the shifted stimuli compared with 2 hr of training with feedback in our study. While both studies reported that performance had saturated by the end of this period, it should be noted that the listeners in the Yoon et al. study only reached 30% correct with shifted stimuli in the +10-dB TMR condition, which meant that the low percent correct scores in the +5-dB TMR condition were likely limited by floor effects. In contrast, our listeners achieved 70% correct in the +10-dB TMR that provided a much larger range of performance for us to explore the effects of shift at lower TMRs. Our explicit training with feedback was likely long enough to saturate performance (Waked et al. 2017), and the use of a small closed set of words in matrix sentences helped us to avoid issues with floor effects.

There are also some other noteworthy, albeit more minor, methodological differences between these studies. First, in Yoon et al. (2013), they tested frequency-to-place mismatches of 5 and 8 mm and interaural frequency mismatches of $\Delta = 0$ and 2 mm (i.e., the right ear had more frequency-to-place mismatch). In contrast, we tested frequency-to-place mismatches of 0, 3, 4.5, and 6 mm and interaural frequency mismatches of Δs as large as ±6 mm in the present study. Therefore, they had no control unshifted conditions and we tested much larger mismatches, 3 mm or greater, which is on the order of shift for which we consider necessary to see degradation to binaural processing (Goupell 2015; Goupell et al. 2013; Kan et al. 2013, 2015).

Next, Yoon et al. (2013) used an envelope cutoff of 160 Hz; in contrast, we used an envelope cutoff of 50 Hz. The reason to have a relatively low envelope cutoff in our study was to limit the frequency range of the sidebands of the sinusoidal carriers that were created when the speech information was modulated by the speech envelopes. Unique to this type of study with frequency shift, relatively small shifts might have resolved sidebands for low-frequency channels where the auditory filters are narrower, but larger basalward shifts might contain only channels with unresolved sidebands where auditory filters are wider. Resolved sidebands are known to vastly improve speech understanding using vocoders with sinusoidal carriers (Souza & Rosen 2009). Both studies seem to potentially have this confound for the lowest frequency channels; however, ours was limited to the two lowest frequency channels and Yoon et al. (2013) was limited to five channels.

Future work to better understand when unmasking and interference occur in CIs may allow the components of SRM to be calculated like in Yoon et al. (2013). However, when conducting this type of work, one should consider the need to use stimuli designed to produce larger amounts of SRM, along with noting the importance of training for shifted or warped stimuli when testing with naive listeners (Rosen et al. 1999). The vocoders that were implemented in Yoon et al. and the present study may have incorrectly estimated the effects of mismatch on SRM because they do not simulate current spread (Oxenham & Kreft 2014). Perhaps the nonmonotonicities observed in speech understanding differences (particularly ±4.5-mm Δ or 1.5-channel shift, Fig. 4A) would not have occurred with more spectral overlap between channels. Furthermore, there are assumptions about where the stimuli are perceived spatially in these studies. Assessing the relationship among perceived spatial separation, fusion, and SRM would be beneficial in future studies.

This work was motivated by the fact that there is potential for mismatches in place of stimulation in bilateral CI users (e.g., Goupell 2015; Kan et al. 2013, 2015; Pearl et al. 2014), and the current results have implications for programming of sound processors in these patients. There is currently no standard clinical procedure for determining how much interaural mismatch occurs for an individual. These data, along with the aforementioned results in bilateral CI users, suggest that clinically mismatched frequency allocation tables in the CI processors would introduce interaural envelope decorrelation, which could severely impact sound localization abilities and speech understanding in noise. The best approach for aligning the clinical stimulation might have two stages. First, there would need to be an estimate of electrode locations, possibly through some combination of examining asymmetries in computed tomography scans (Long et al. 2014; Noble et al. 2011), perceptual tasks such as ITD discrimination or pitch matching tasks (Kan et al. 2013, 2015; Long et al. 2003; Poon et al. 2009), or electrophysiology (Hu & Dietz 2015). Second, interaurally matched electrodes should have the same frequency allocation. Even then, interaural differences in neural survival could limit the ability to interaurally match the inputs.

One question that remains is whether neural plasticity could potentially rectify the issue of interaural mismatch. CI users likely experience some frequency-to-place mismatch given the standard intracochlear placement of the electrode arrays and default frequency allocation tables used to program the devices. CI users also improve in understanding speech after months to years of frequency-to-place mismatch (Blamey et al. 2013), which may be partially a result of continual adjustments in the clinic to optimize performance, and partially a result of neural plasticity and relearning the encoding of different speech phonemes. While plasticity may occur for speech understanding (Blamey et al. 2013) and other basic percepts like pitch (Reiss et al. 2014, 2015), there seems to be little evidence suggesting that there is plasticity that would correct for interaural frequency mismatch in the binaural circuits of the brainstem. In fact, Fallon et al. (2014) have argued that plasticity in cats with CIs is more likely to occur at the level of central auditory circuits (such as primary auditory cortex) than at the level of the brainstem or midbrain. If the binaural circuits at the level of the brainstem and midbrain cannot compensate for interaural frequency mismatch, it would be critical to address this problem through clinical mapping. This might be particularly important for young children who are implanted bilaterally at a young age, as results to date suggest that early bilateral stimulation does not guarantee that children are good at sound localization or binaural unmasking of speech (for a review, see Litovsky & Gordon 2016).

CONCLUSION

Frequency-to-place mismatch and interaural frequency mismatch were simulated using an eight-channel vocoder. Speech understanding and SRM were measured. Increasing bilateral shifts (frequency-to-place mismatch) produced larger decreases in speech understanding but did not affect SRM. Increasing unilateral shifts (interaural frequency mismatch) produced larger decreases in speech understanding, and significantly decreased SRM when the ear with the better TMR was shifted. If bilateral CIs are to provide spatial hearing benefits like the large amounts of SRM in NH listeners, it would be necessary to compensate for interaural frequency mismatch.

REFERENCES

Aguiar, D. E., Taylor, N. E., Li, J., et al. (2016). Information theoretic evaluation of a noiseband-based cochlear implant simulator. *Hear Res.*, *333*, 185–193.

Algazi, V. R., Avendano, C., Duda, R. O. (2001). Elevation localization and head-related transfer function analysis at low frequencies. *J Acoust Soc Am*, *109*, 1110–1122.

Başkent, D., & Shannon, R. V. (2007). Combined effects of frequency compression-expansion and shift on speech recognition. *Ear Hear*, *28*, 277–289.

Bernstein, L. R., & Trahiotis, C. (1996). On the use of the normalized correlation as an index of interaural envelope correlation. *J Acoust Soc Am*, *100*, 1754–1763.

Bernstein, J. G. W., Goupell, M. J., Schuchman, G. I., et al. (2016). Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees. *Ear Hear*, *37*, 282–288.

Best, V., Thompson, E. R., Mason, C. R., et al. (2013). An energetic limit on spatial release from masking. *J Assoc Res Otolaryngol*, *14*, 603–610.

Blamey, P., Artieres, F., Başkent, D., et al. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiol Neurootol*, *18*, 36–47.

Blanks, D. A., Buss, E., Grose, J. H., et al. (2008). Interaural time discrimination of envelopes carried on high-frequency tones as a function of level and interaural carrier mismatch. *Ear Hear*, *29*, 674–683.

Blanks, D. A., Roberts, J. M., Buss, E., et al. (2007). Neural and behavioral sensitivity to interaural time differences using amplitude modulated tones with mismatched carrier frequencies. *J Assoc Res Otolaryngol*, *8*, 393–408.

Blauert, J., & Lindemann, W. (1986). Auditory spaciousness: Some further psychoacoustic analyses. *J Acoust Soc Am*, *80*, 533–542.

Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J Acoust Soc Am*, *83*, 1508–1516.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am*, *109*, 1101–1109.

Brungart, D. S., Simpson, B. D., Ericson, M. A., et al. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am*, *110*(), 2527–2538.

Buss, E., Pillsbury, H. C., Buchman, C. A., et al. (2008). Multicenter U.S. bilateral MED-EL cochlear implantation study: Speech perception over the first year of use. *Ear Hear*, *29*, 20–32.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am*, *25*, 975–979.

Darwin, C. J., & Hukin, R. W. (2000). Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *J Acoust Soc Am*, *107*, 970–977.

Dorman, M. F., Loizou, P. C., Rainey, D. (1997). Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. *J Acoust Soc Am*, *102*(), 2993–2996.

Durlach, N. I., Mason, C. R., Kidd, G., Jr, et al. (2003). Note on informational masking. *J Acoust Soc Am*, *113*, 2984–2987.

Fallon, J. B., Shepherd, R. K., Irvine, D. R. (2014). Effects of chronic cochlear electrical stimulation after an extended period of profound deafness on primary auditory cortex organization in cats. *Eur J Neurosci*, *39*, 811–820.

Feddersen, W. E., Sandel, T. T., Teas, D. C., et al. (1957). Localization of high-frequency tones. *J Acoust Soc Am*, *29*, 988–991.

Francart, T., & Wouters, J. (2007). Perception of across-frequency interaural level differences. *J Acoust Soc Am*, *122*, 2826–2831.

Freyman, R. L., Balakrishnan, U., Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *J Acoust Soc Am*, *109*(), 2112–2122.

Freyman, R. L., Balakrishnan, U., Helfer, K. S. (2008). Spatial release from masking with noise-vocoded speech. *J Acoust Soc Am*, *124*, 1627–1637.

Fu, Q. J., & Galvin, J. J. 3rd. (2003). The effects of short-term training for spectrally mismatched noise-band speech. *J Acoust Soc Am*, *113*, 1065–1072.

Fu, Q. J., & Shannon, R. V. (1999). Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *J Acoust Soc Am*, *105*, 1889–1900.

Fu, Q. J., Shannon, R. V., Galvin, J. J. 3rd. (2002). Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant. *J Acoust Soc Am*, *112*, 1664–1674.

Gabriel, K. J., & Colburn, H. S. (1981). Interaural correlation discrimination: I. Bandwidth and level dependence. *J Acoust Soc Am*, *69*, 1394–1401.

Goupell, M. J. (2010). Interaural fluctuations and the detection of interaural incoherence. IV. The effect of compression on stimulus statistics. *J Acoust Soc Am*, *128*, 3691–3702.

Goupell, M. J. (2015). Interaural correlation-change discrimination in bilateral cochlear-implant users: Effects of interaural frequency mismatch, centering, and age of onset of deafness. *J Acoust Soc Am*, *137*, 1282–1297.

Goupell, M. J., & Litovsky, R. Y. (2014). The effect of interaural fluctuation rate on correlation change discrimination. *J Assoc Res Otolaryngol*, *15*, 115–129.

Goupell, M. J., Kan, A., Litovsky, R. Y. (2016). Spatial attention in bilateral cochlear-implant users. *J Acoust Soc Am*, *140*, 1652–1662.

Goupell, M. J., Stakhovskaya, O. A., Bernstein, J. G. W. (2017). Contralateral interference caused by binaurally presented competing speech in adult bilateral cochlear-implant users. *Ear Hear*, *39*, 110–123.

Goupell, M. J., Stoelb, C., Kan, A., et al. (2013). Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening. *J Acoust Soc Am*, *133*, 2272–2287.

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

Gstoettner, W., Franz, P., Hamzavi, J., et al. (1999). Intracochlear position of cochlear implant electrodes. *Acta Otolaryngol*, *119*, 229–233.

Hallberg, L. R., Hallberg, U., Kramer, S. E. (2008). Self-reported hearing difficulties, communication strategies and psychological general well-being (quality of life) in patients with acquired hearing impairment. *Disabil Rehabil*, *30*, 203–212.

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. *J Acoust Soc Am*, *115*, 833–843.

Helbig, S., Mack, M., Schell, B., et al. (2012). Scalar localization by computed tomography of cochlear implant electrode carriers designed for deep insertion. *Otol Neurotol*, *33*, 745–750.

Henning, G. B. (1974). Detectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am*, *55*, 84–90.

Hu, H., Dietz, M. (2015). Comparison of interaural electrode pairing methods for bilateral cochlear implants. *Trends Hear*, *19*, 1–22.

Kan, A., & Litovsky, R. Y. (2015). Binaural hearing with electrical stimulation. *Hear Res*, *322*, 127–137.

Kan, A., Litovsky, R. Y., Goupell, M. J. (2015). Effects of interaural pitch matching and auditory image centering on binaural sensitivity in cochlear implant users. *Ear Hear*, *36*, e62–e68.

Kan, A., Stoelb, C., Litovsky, R. Y., et al. (2013). Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant users. *J Acoust Soc Am*, *134*, 2923–2936.

Ketten, D. R., Skinner, M. W., Wang, G., et al. (1998). In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays. *Ann Otol Rhinol Laryngol Suppl*, *175*, 1–16.

Kidd, G., Jr, Best, V., Mason, C. R. (2008). Listening to every other word: Examining the strength of linkage variables in forming streams of speech. *J Acoust Soc Am*, *124*, 3793–3802.

Kramer, S. E., Kapteyn, T. S., Festen, J. M. (1998). The self-reported handicapping effect of hearing disabilities. *Audiology*, *37*, 302–312.

Laback, B., Pok, S. M., Baumgartner, W. D., et al. (2004). Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. *Ear Hear*, *25*, 488–500.

Landsberger, D. M., Svrakic, M., Roland, J. T., Jr, et al. (2015). The relationship between insertion angles, default frequency allocations, and spiral ganglion place pitch in cochlear implants. *Ear Hear*, *36*, e207–e213.

Litovsky, R. Y., & Gordon, K. (2016). Bilateral cochlear implants in children: Effects of auditory experience and deprivation on auditory perception. *Hear Res*, *338*, 76–87.

Litovsky, R. Y., Goupell, M. J., Godar, S., et al. (2012). Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory. *J Am Acad Audiol*, *23*, 476–494.

Litovsky, R., Parkinson, A., Arcaroli, J., et al. (2006). Simultaneous bilateral cochlear implantation in adults: A multicenter clinical study. *Ear Hear*, *27*, 714–731.

Loizou, P. C., Hu, Y., Litovsky, R., et al. (2009). Speech recognition by bilateral cochlear implant users in a cocktail-party setting. *J Acoust Soc Am*, *125*, 372–383.

Loizou, P. C. (2006). Speech processing in vocoder-centric cochlear implants. In A. Moller (Ed.), *Cochlear and Brainstem Implants*, (pp. 109–143).Basel, Switzerland: Karger.

Long, C. J., Eddington, D. K., Colburn, H. S., et al. (2003). Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user. *J Acoust Soc Am*, *114*, 1565–1574.

Long, C. J., Holden, T. A., McClelland, G. H., et al. (2014). Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding. *J Assoc Res Otolaryngol*, *15*, 293–304.

Misurelli, S. M., & Litovsky, R. Y. (2015). Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity. *J Acoust Soc Am*, *138*, 319–331.

Mueller, J., Schoen, F., Helms, J. (2002). Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system. *Ear Hear*, *23*, 198–206.

Nelson, D. A., Donaldson, G. S., Kreft, H. (2008). Forward-masked spatial tuning curves in cochlear implant users. *J Acoust Soc Am*, *123*, 1522–1543.

Nelson, D. A., Kreft, H. A., Anderson, E. S., et al. (2011). Spatial tuning curves from apical, middle, and basal electrodes in cochlear implant users. *J Acoust Soc Am*, *129*, 3916–3933.

Noble, J. H., Labadie, R. F., Majdani, O., et al. (2011). Automatic segmentation of intracochlear anatomy in conventional CT. *IEEE Trans Biomed Eng*, *58*, 2625–2632.

Nuetzel, J. M., Hafter, E. R. (1981). Discrimination of interaural delays in complex waveforms: Spectral effects. *J Acoust Soc Am*, *69*, 1112–1118.

Oxenham, A. J., Kreft, H. A. (2014). Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing. *Trends Hear*, *18*, 1–14.

Pearl, M. S., Roy, A., Limb, C. J. (2014). High-resolution secondary reconstructions with the use of flat panel CT in the clinical assessment of patients with cochlear implants. *AJNR Am J Neuroradiol*, *35*, 1202–1208.

Peters, B. R., Wyss, J., Manrique, M. (2010). Worldwide trends in bilateral cochlear implantation. *Laryngoscope*, *120*(Suppl 2), S17–S44.

Poon, B. B., Eddington, D. K., Noel, V., et al. (2009). Sensitivity to interaural time difference with bilateral cochlear implants: Development over time and effect of interaural electrode spacing. *J Acoust Soc Am*, *126*, 806–815.

Reiss, L. A., Ito, R. A., Eggleston, J. L., et al. (2015). Pitch adaptation patterns in bimodal cochlear implant users: Over time and after experience. *Ear Hear*, *36*, e23–e34.

Reiss, L. A., Turner, C. W., Karsten, S. A., et al. (2014). Plasticity in human pitch perception induced by tonotopically mismatched electro-acoustic stimulation. *Neuroscience*, *256*, 43–52.

Rosen, S., Faulkner, A., Wilkinson, L. (1999). Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. *J Acoust Soc Am*, *106*, 3629–3636.

Rothauser, E. H., Chapman, W., Guttman, N., et al. (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans Acoust Speech Signal Proc*, *17*, 225–246.

Schleich, P., Nopp, P., D'Haese, P. (2004). Head shadow, squelch, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant. *Ear Hear*, *25*, 197–204.

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

Shannon, R. V., Zeng, F. G., Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. *J Acoust Soc Am*, *104*, 2467–2476.

Siciliano, C. M., Faulkner, A., Rosen, S., et al. (2010). Resistance to learning binaurally mismatched frequency-to-place maps: Implications for bilateral stimulation with cochlear implants. *J Acoust Soc Am*, *127*, 1645–1660.

Skinner, M. W., Holden, L. K., Holden, T. A. (1995). Effect of frequency boundary assignment on speech recognition with the speak speech-coding strategy. *Ann Otol Rhinol Laryngol Suppl*, *166*, 307–311.

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.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J Speech Hear Res*, *28*, 455–462.

Svirsky, M. A., Silveira, A., Neuburger, H., et al. (2004). Long-term auditory adaptation to a modified peripheral frequency map. *Acta Otolaryngol*, *124*, 381–386.

van Besouw, R. M., Forrester, L., Crowe, N. D., et al. (2013). Simulating the effect of interaural mismatch in the insertion depth of bilateral cochlear implants on speech perception. *J Acoust Soc Am*, *134*, 1348–1357.

van Hoesel, R. J., & Clark, G. M. (1997). Psychophysical studies with two binaural cochlear implant subjects. *J Acoust Soc Am*, *102*, 495–507.

van Hoesel, R. J., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *J Acoust Soc Am*, *113*, 1617–1630.

van Hoesel, R. J. M., Böhm, M., Pesch, J., et al. (2008). Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies. *J Acoust Soc Am*, *123*, 2249–2263.

Waked, A., Dougherty, S., Goupell, M. J. (2017). Vocoded speech understanding with simulated shallow insertion depths in adults and children. *J Acoust Soc Am*, *141*, EL45–EL50.

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

Whitmer, W. M., Seeber, B. U., Akeroyd, M. A. (2012). Apparent auditory source width insensitivity in older hearing-impaired individuals. *J Acoust Soc Am*, *132*, 369–379.

Yoon, Y. S., Liu, A., Fu, Q. J. (2011). Binaural benefit for speech recognition with spectral mismatch across ears in simulated electric hearing. *J Acoust Soc Am*, *130*, EL94–E100.

Yoon, Y. S., Shin, Y. R., Fu, Q. J. (2013). Binaural benefit with and without a bilateral spectral mismatch in acoustic simulations of cochlear implant processing. *Ear Hear*, *34*, 273–279.