



Published in final edited form as:

Ear Hear. 2016 ; 37(3): 289–302. doi:10.1097/AUD.0000000000000284.

Having Two Ears Facilitates the Perceptual Separation of Concurrent Talkers for Bilateral and Single-Sided Deaf Cochlear Implantees

Joshua G. W. Bernstein¹, Matthew J. Goupell², Gerald I. Schuchman¹, Arnaldo L. Rivera^{3,4}, and Douglas S. Brungart¹

¹ National Military Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, Bethesda, Maryland, USA.

² Department of Hearing and Speech Sciences, University of Maryland – College Park, Maryland, USA.

³ Department of Otolaryngology Head and Neck Surgery, Walter Reed National Military Medical Center, Bethesda, Maryland, USA.

Abstract

Objectives—Listening to speech with multiple competing talkers requires the perceptual separation of the target voice from the interfering background. Normal-hearing (NH) listeners are able to take advantage of perceived differences in the spatial locations of competing sound sources to facilitate this process. Previous research suggests that bilateral (BI) cochlear-implant (CI) listeners cannot do so, and it is unknown whether single-sided deaf CI users (SSD-CI; one acoustic and one CI ear) have this ability. This study investigated whether providing a second ear via cochlear implantation can facilitate the perceptual separation of targets and interferers in a listening situation involving multiple competing talkers.

Design—BI-CI and SSD-CI listeners were required to identify speech from a target talker mixed with one or two interfering talkers. In the baseline monaural condition, the target speech and the interferers were presented to one of the CIs (for the BI-CI listeners) or to the acoustic ear (for the SSD-CI listeners). In the bilateral condition, the target was still presented to the first ear but the interferers were presented to both the target ear and the listener's second ear (always a CI), thereby testing whether CI listeners could use information about the interferer obtained from a second ear to facilitate perceptual separation of the target and interferer.

Address for correspondence: Joshua G. W. Bernstein, National Military Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, 4954 Palmer Rd., Bethesda, MD 20889, USA, Telephone: 301-319-7040, Fax: 301-319-7076, joshua.g.bernstein.civ@mail.mil.

⁴Current affiliation: Department of Otolaryngology – Head and Neck Surgery, University of Missouri School of Medicine, Columbia, Missouri, USA.

Portions of this article were presented at the 2013 and 2015 Conferences on Implantable Auditory Prostheses, Tahoe City, California and at the 2015 Midwinter Meeting of the Association for Research in Otolaryngology, Baltimore, Maryland.

The views expressed in this article are those of the authors and do not reflect the official policy of the Department of the Army, the Department of the Navy, the Department of the Air Force, the Department of Defense, or the U.S. Government.

Conflicts of Interest:

For the remaining authors, none were declared.

Results—Presenting a copy of the interfering signals to the second ear improved performance, up to 4-5 dB (12-18 percentage points), but the amount of improvement depended on the type of interferer. For BI-CI listeners, the improvement occurred mainly in conditions involving one interfering talker, regardless of gender. For SSD-CI listeners, the improvement occurred in conditions involving one or two interfering talkers of the same gender as the target. This interaction is consistent with the idea that the SSD-CI listeners had access to pitch cues in their NH ear to separate the opposite-gender target and interferers, while the BI-CI listeners did not.

Conclusions—These results suggest that a second auditory input via a CI can facilitate the perceptual separation of competing talkers in situations where monaural cues are insufficient to do so, thus partially restoring a key advantage of having two ears that was previously thought to be inaccessible to CI users.

Keywords

Cochlear implant; Binaural; Auditory scene analysis; Informational masking

INTRODUCTION

People are often surrounded by multiple spatially separated sound sources that compete with one another for the listener's attention. This is particularly true in environments like restaurants, cocktail parties or conference rooms, where individuals trying to communicate must compete with many other simultaneous conversations. In these situations, the listener must parse the complex auditory scene to focus on the sound source of interest (the “target”) while ignoring the background sounds (the “interferers”). When the target and interferers are perceptually dissimilar (e.g., speech and noise), performance is limited mainly by energetic masking, whereby energy from the interferer overwhelms the target and renders portions of it undetectable. However, situations involving multiple competing voices add an additional dimension of difficulty to the task, because the target and interferers are acoustically similar to one another. This can impede the listener's ability to perceptually separate concurrent sound sources even if there is relatively little overlap in time and frequency (for a review, see Kidd et al. 2008).

For normal-hearing (NH) listeners, several auditory cues are available to facilitate the perceptual separation of target and interfering speech. Listeners can take advantage of voice pitch or timbre differences to differentiate between voices (Darwin & Hukin 2000). Listeners can also take advantages of differences in spatial locations (Arbogast et al. 2002). Spatial cues become all the more important in situations involving multiple voices with similar qualities – such as when the target and interferers are the same gender – because pitch and timbre become less reliable cues to differentiate the concurrent voices (Darwin et al. 2003). In these situations, having two ears provides a tremendous advantage by allowing a listener to make use of spatial differences to perceptually separate the target and interferers (e.g., Marrone et al. 2008; Martin et al. 2012). As a result, individuals with only one functional ear are at a distinct disadvantage in complex listening environments compared to those with two functional ears.

The benefits associated with having two ears have led to changes in the use of cochlear implants (CIs) to treat deafness. CIs deliver an impoverished representation of the acoustic signal to the auditory nerve, consisting of slow modulations (typically <400 Hz; Loizou 2006) delivered to an electrode array positioned along the length of the cochlea. Despite the crudeness of the electrical representation in several dimensions, including poor spectral resolution (Nelson et al. 2008) and a lack of strong pitch information (Chatterjee & Peng 2008), many CI listeners understand speech quite well, especially in quiet conditions (Gifford et al. 2008). While historically, deaf individuals received only one CI, in recent years most pre-lingually deafened children, and approximately 10% of post-lingually deafened adults, receive two CIs in an attempt to provide some of the two-ear benefits that are experienced by NH listeners (Peters et al. 2010). Even more recently, a small number of individuals with single-sided deafness (SSD; one normal ear and one deaf ear) have received a CI in their deaf ear, thereby providing them with two independent auditory inputs (e.g., Vermeire & van de Heyning 2009; Arndt et al. 2011; Firszt et al. 2012; Hansen et al. 2013; Erbele et al. 2015; Zeitler et al. 2015)¹.

Previous literature for bilateral CI (BI-CI) listeners and the limited existing literature for SSD-CI listeners have generally shown that having two auditory inputs can provide moderate benefits for sound localization and for speech understanding in noise when there is a spatial separation between target and masker (e.g., Vermeire & van de Heyning, 2009; Arndt et al. 2011; Litovsky et al. 2012; Tokita et al. 2014; Zeitler et al. 2015). However, having two ears provides these listeners with a more limited set of advantages than those experienced by NH listeners. For NH listeners, the advantages of having two ears for listening to speech in noisy situations can be broadly characterized in two categories (Zurek 1993). The first category is a “better-ear benefit”, which arises because the acoustic shadow of the head attenuates sounds that arrive at the ear on the opposite side of the head from the sound source. This head-shadowing effect results in a difference in signal-to-noise ratio (SNR) between the ears whenever the target and interferers originate from different spatial locations. Listeners with two ears can take advantage of this difference by selectively attending to the ear with the highest SNR. The effective increase in SNR obtained by using this strategy is referred to as the better-ear benefit. Although this benefit can exceed 20 dB at high frequencies in an anechoic environment (Bronkhorst & Plomp 1988), the maximum better-ear benefit for speech in a reverberant environment is on the order of 5-8 dB (Culling et al. 2012). The second category of two-ear advantage, referred to as “binaural interaction,” requires the auditory system to calculate differences between the signals arriving at the two ears. A sound source originating on one side of the head arrives at the closer ear both earlier and at a higher intensity than at the farther ear. The NH auditory system is able to utilize the different interaural time differences (ITDs) and the interaural level differences (ILDs) of spatially separated target and interfering sounds to help improve the detection of the target signal. This effect can theoretically improve the detection threshold for a target signal that is spatially separated from a noise interferer by as much as 15 dB in addition to any better-ear benefit (Wan et al. 2010). However, in real-world listening environments with spatially separated interfering noises, an improvement of 3-4 dB due to binaural interactions is more

¹Cochlear implants are not current labeled by the United States Food & Drug Administration for use for the treatment of SSD.

typical (Hawley et al. 2004). In cases where the interfering sounds in the environment are potentially confusable with the target sound, binaural interactions can provide an even larger advantage by facilitating the perceptual separation of target and interfering sounds based on differences in their apparent locations (Hawley et al. 2004; Gallun et al. 2005).

For BI-CI listeners, the preponderance of evidence suggests that the observed two-ear benefit for speech perception in the presence of interfering sounds primarily reflects a better-ear advantage. Many studies show little evidence of binaural-interaction effects or improved perceptual source separation (e.g., van Hoesel et al. 2008; Litovsky et al., 2009; Loizou et al. 2009; Reeder et al., 2014), although data from Eapen et al. (2009) suggest that a binaural-interaction advantage can arise after a year or more of bilateral listening experience. For SSD-CI listeners, two-ear benefits have only been observed for spatial configurations where the CI ear has the more favorable SNR, suggesting that the CI is providing, for the most part, a better-ear benefit (Vermeire & van de Heyning 2009; Arndt et al. 2011; Zeitler et al. 2015). Numerous psychophysical and physiological studies have established that the binaural system is sensitive to ITDs and ILDs when well-controlled stimuli are delivered directly to single electrodes in each ear, bypassing the complex signal processing associated with the external sound processor (e.g., Lu et al. 2010; Hancock et al. 2012; Chung et al. 2014; Goupell & Litovsky 2015). BI-CI listeners are able to make use of these interaural differences to discriminate small changes in sound location (e.g., Kan et al. 2013) and to improve the detection of simple electrical pulse trains presented in noise (e.g., Lu et al. 2010; Goupell & Litovsky 2015). For speech stimuli, there are at least two possible reasons why previous studies have typically shown that BI-CI and SSD-CI listeners show very little performance benefit from binaural-interaction effects or from the perceptual separation of spatially separated sound sources. The first is that the interaural-difference cues that are critical to binaural processing and spatial perception are either absent from or poorly encoded in the CI signal for complex sound sources presented in the free field and processed by the external sound processor (e.g., Dorman et al. 2014). The second is that the previous studies have used target and interfering sounds that were not confusable enough to require perceptual separation on the basis of apparent spatial location in order to be segregated from the target speech.

Indeed, Bernstein et al. (2015) did find some evidence that simulated SSD-CI listeners could obtain a perceptual-separation advantage in an experimental paradigm that simultaneously addressed both of these issues. First, they replaced the ITD and ILD cues that would occur for typical sound sources in the free field with the extreme case where the target speech was presented monaurally to the acoustic ear and the interfering talkers were presented to both the normal ear and to a simulated CI ear using an eight-channel vocoder. Vocoder processing simulates aspects of CI speech processing by passing auditory signals through a bank of bandpass filters and extracting acoustic envelope information. In lieu of electrical stimulation, these envelopes are delivered acoustically to the NH cochlea by modulating narrowband noise carriers (Shannon et al. 1995; Loizou 2006). For NH listeners presented with unprocessed speech, this configuration produces a very salient difference in the apparent spatial locations of the target and masker without relying on natural ITD and ILD cues that are poorly encoded by CIs. This paradigm ensured that any increase in performance from adding the second vocoded ear could be attributed to improved perceptual

source separation, and not to a better-ear benefit, because the first ear was the only one that contained the target and therefore was always the better ear. Second, Bernstein et al. (2015) selected speech materials that allowed systematic manipulation of the perceptual similarity of the target and masking signal (Brungart 2001). The study found that presenting the vocoded interferers to the second ear improved performance, but only in the conditions involving same-gender interfering talkers, where listeners likely experienced difficulty perceptually separating the sources monaurally due to a relative lack of pitch and timbre differences between the target and interferers.

We hypothesized that BI-CI and SSD-CI listeners would also demonstrate the ability to perceptually separate competing target and interfering speech using this paradigm. Performance in the monaural condition (target and interferers presented to one ear) was compared to performance in the bilateral condition (target presented to one ear and the interferers presented to both ears) to determine whether the presentation of the interferers to both ears would facilitate the perceptual separation of the concurrent talkers. The number and gender of the interfering talkers was manipulated to alter the difficulty the listeners would experience in perceptually separating the concurrent speech streams. A stationary-noise interferer was included as a control condition where listeners were expected to experience little difficulty in perceptually separating the target and interferer based on their distinct acoustic differences. Previous results have shown that the difficulty in separating concurrent speech sounds (Brungart, 2001) and the magnitude of spatial release from masking (Freyman et al., 2008) can vary considerably as a function of the target-to-masker ratio (TMR). Therefore, rather than use an adaptive test that adjusts performance to a desired percentage-correct level of performance, a method of fixed stimuli was used to measure performance, with a wide range of TMRs tested in each condition.

In addition to a group of BI-CI listeners and a group of SSD-CI listeners that participated in the study, a group of NH listeners were also included as a control. The NH listeners were presented with unprocessed stimuli, and with noise-vocoded CI simulations to estimate the possible upper limit in the amount of perceptual source-separation advantage that might be experienced by a BI-CI or SSD-CI listener in this task. Vocoder processing simulates some aspects of CI processing (most CI processing algorithms are vocoder-centric; Loizou 2006) while greatly reducing the intersubject variability in the CI population likely attributable to differences in duration of deafness, neural survival, plasticity, surgical outcome, and CI programming parameters (Blamey et al. 2013).

METHODS

Listeners

Three groups of listeners participated in the study: BI-CI listeners (two CIs), SSD-CI listeners (one acoustic-hearing ear and a CI in the deaf ear), and NH listeners. The study procedures were approved by the Institutional Review Boards at both participating institutions.

BI-CI listeners—Nine BI-CI listeners participated. Demographic information for these listeners is provided in Table 1. All BI-CI listeners were tested at the University of Maryland

– College Park. Eight of the BI-CI listeners were post-lingually deafened, while one was pre- or peri-lingually deafened. All nine BI-CI listeners were implanted sequentially as adults with Cochlear Ltd. devices.

SSD-CI listeners—Seven individuals with SSD participated. Demographic information for these listeners is provided in Table 2. All SSD-CI listeners were tested at Walter Reed National Military Medical Center. All lost hearing in their deaf ear as adults, and were implanted with MED-EL or Cochlear Ltd. devices. Three of the listeners had NH in the acoustic ear. Two of the listeners had mild sensorineural hearing loss. One listener had a mixed loss, with severe high-frequency sensorineural hearing loss (above 3000 Hz) and a mild conductive loss. One listener had a conductive loss, but normal cochlear function (normal bone-conduction audiogram). Air-conduction audiograms for the seven SSD-CI listeners (and the bone-conduction audiogram for the two listeners with conductive hearing loss) are provided in Table 3.

NH listeners—Eight individuals with bilaterally NH (six female, age range 21-36 years) participated. All NH listeners were tested at the University of Maryland – College Park. They had normal air-conduction audiometric thresholds 15 dB hearing level (HL) in both ears at octave frequencies between 250 and 4000 Hz, 25 dB HL at 8000 Hz, and had no interaural asymmetries greater than 10 dB.

Procedure

The call sign-based word-identification task based on the coordinate response measure (CRM; Brungart 2001) was employed because the identical sentence structures of the target and interfering sentences make it difficult to perceptually separate the concurrent voices. The difficulty of the task was varied by adjusting the number (one or two) and gender (same or opposite) of the interfering talkers, or presenting the same stimuli in a speech-shaped stationary noise. CRM sentences are of the form “Ready (call sign) go to (color) (number) now” spoken by one of four different male or four different female talkers. Listeners were instructed to follow the speech of the talker who used the call sign “Baron,” and identify the color and number spoken by this target talker, while ignoring the one or two interferers that used other call signs (e.g., “Eagle” or “Arrow”). The response matrix consisted of a four-by-eight array of virtual buttons, with each row representing a given color and each column representing a given number. On each trial, the target stimulus was randomly selected from among the eight possible talkers, eight possible numbers (one through eight) and four possible colors (red, green, white, and blue).

Five different interferer types were tested. Four of the interferer types involved interfering talkers (one or two interfering talkers, with both interferers either the same or opposite gender as the target talker). The target and interferers were constrained such that each signal was spoken by a different individual, with a different number and color. The fifth interferer type was speech-shaped stationary noise. On each trial, a noise was generated with the spectrum of speech produced by a talker of the same gender as the target. This was done to produce similar spectra for the target speech and the noise. A sentence spoken by a single same-gender interfering talker was selected, and the fast-Fourier transform (FFT) of the

signal was computed. The phases of the FFT were then randomized, and the inverse FFT was taken to generate the stationary speech-spectrum shaped noise.

For all interferer types, the target and interferer signals were combined at the desired TMR and presented to the first ear (referred to as the target ear). For the other (non-target) ear, two stimulus configurations were tested. In the bilateral configuration, the same interferer signals were also presented to the second ear but without the target. In the monaural configuration, no stimulus was presented to the non-target ear. The perceptual source-separation advantage was evaluated by comparing performance for the monaural and bilateral configurations.

The NH listeners were tested in three signal-processing conditions. In the unprocessed condition, the signals presented to both ears remained unprocessed as a control condition to characterize the perceptual-separation benefit for individuals with two NH ears. In the other two signal-processing conditions, the signals presented to one or both ears were processed with an eight-channel noise-band vocoder after the target and interferer stimuli were combined and adjusted in level. For the bilateral-vocoder condition, the signals presented to both ears were processed by noise vocoders, with independent noise carriers in the two ears. For the SSD-vocoder condition, the combined target and interferer stimuli were presented unprocessed to the left ear, while the stimulus consisting of only the interferers was vocoded before being presented to the right ear.

Vocoder processing was carried out as described by Hopkins et al. (2008) and Bernstein et al. (2015). First, the acoustic signals were passed through a finite-impulse response filterbank that separated the signal into eight frequency channels. The filterbank covered a frequency range of 100-10,000 Hz, had bandwidths proportional to the equivalent rectangular bandwidth of a NH auditory filter (Glasberg & Moore 1990), and had slopes that varied slightly but were at least 100 dB/octave for all channels. Second, the Hilbert amplitude envelope was extracted from the signal in each channel. Third, the envelopes were used to modulate independent white noise carriers, with the resulting signals passed through the original filterbank before normalizing the level in each channel to equal the root-mean-squared level of the original unprocessed signal in that band. Finally, the signals were summed across channels to generate the broadband vocoded sound.

Stimuli were delivered using custom MATLAB software via a circumaural headphone (for an acoustic ear), or via direct connection to the auxiliary input of the sound processor (for a CI ear). At University of Maryland-College Park, stimuli were presented using an Edirol (Roland Corporation, Hamamatsu, Japan) soundcard to the CI auxiliary input for the BI-CI listeners or HD650 circumaural headphones (Sennheiser, Wedemark, Germany) for the NH listeners. At Walter Reed National Military Medical Center (SSD-CI listeners), stimuli were presented using Tucker-Davis (Alachua, FL) System III hardware (an enhanced real-time processor, TDT RP2.1, for D/A conversion and a headphone buffer, TDT HB7), closed HD280 circumaural headphones (Sennheiser, Wedemark, Germany) for the acoustic-hearing ear and an auxiliary cable for the CI ear. For the SSD-CI listeners, the target ear was always the acoustic-hearing ear. For the NH listeners, the target ear was always the left ear. All nine BI-CI listeners completed the experiment with the target stimulus presented to the left ear.

Additionally, six of the BI-CI listeners completed the experiment with the target stimulus presented to the right ear.

Each listener was tested over a range of five or six TMRs in each condition. Table 4 lists the TMRs tested for each listener group and condition. Slightly different ranges were employed for the different interferer types for the NH and SSD-CI listeners, due to the expectation that conditions involving a single interfering talker would yield a higher level of performance. There was nevertheless a large degree of overlap in the ranges of TMRs tested across the listener groups and test conditions. Trials were blocked, with 35-56 trials presented per block for a given interferer type, stimulus configuration, and signal-processing condition (but with variable TMR). Blocks were presented in pseudo-random order, with one block completed for all combinations of interferer type, stimulus configuration, and signal-processing conditions before a second block was presented for any of the combinations. Each listener was presented with at least 21 trials for each combination of interferer type, stimulus configuration, ear of presentation (for the BI-CI listeners tested in both ear), signal-processing condition (for the NH listeners), and TMR, for a total of at least 30 blocks, with four exceptions. One SSD-CI listener completed only a minimum of 9 trials per condition (15 blocks total). Three BI-CI listeners completed only a minimum of 16 trials per condition (21 blocks total) for each ear configuration, but this was repeated for both ear configurations, such that there were at least 32 trials per condition and a total of 42 blocks across the two configurations.

A reference stimulus level was established for each ear of each listener. The target was presented at this fixed reference level, and the level of the interferer(s) was adjusted to yield the desired TMR. The interferer stimuli were presented to both ears at the same level relative to each ear's reference. For acoustic ears (NH and SSD listeners), the reference stimulus level was 60 dB sound-pressure level (SPL) (unweighted), except for one SSD-CI listener with a large degree of conductive hearing loss (SSD5; Table 3) for whom the reference was 92 dB SPL. While this high stimulus level did not completely compensate for this listener's reduced audibility, it was the maximum level that could be achieved by the system for the range of TMRs tested before peak-clipping occurred. For the SSD-CI listeners, the reference level for the CI ear was established by asking listeners to adjust the level of a series of sample CRM stimuli to match the loudness of the 60-dB (or 92-dB) SPL acoustic speech stimulus. For the BI-CI listeners, the reference level was determined independently for each ear by having the listener adjust the level of a series of target stimuli to their most comfortable level.

Analysis

Repeated-measures binomial logistic regression analyses tested for the presence of a significant perceptual-separation advantage, with separate analyses conducted for the BI-CI and SSD-CI listener groups and for the three signal-processing conditions for the NH listeners. Significant three-way interactions between condition (i.e., monaural or bilateral), TMR, and interferer type were observed for all five analyses. Planned comparisons examined performance differences between the monaural and bilateral conditions for each subject group, signal-processing condition, and interferer type, with the data averaged across

TMRs. No corrections to p-values were applied for these planned comparisons designed to test for the advantage hypothesized (based on the vocoder results of Bernstein et al. 2015) to occur for conditions involving substantial difficulty in perceptually separating the target and interferers in the monaural case. Post-hoc binomial tests compared bilateral and monaural performance at each TMR, with Bonferroni corrections to adjust for the number of TMRs tested for each interferer type, listener group, and (in the case of the NH listeners) signal-processing condition. Data for all of the listeners were included in these analyses.

For the six BI-CI listeners that were tested with the target presented to the right ear in addition to the conditions with the target presented to the left ear, two repeated-measures logistic regression analyses tested for effects of ear of presentation (better-performing vs. worse-performing ear). In one analysis, the better ear was defined for each listener as the ear that yielded the best mean percentage-correct performance across all monaural conditions. However, van Hoesel and Litovsky (2011) argued that using the outcome measure as the means of identifying the better ear can lead to a bias toward observing an effect of ear of presentation when one is not present. Therefore, a second analysis was conducted that defined the better ear based on patient self-report. Neither analysis found a significant main effect of ear of presentation ($p=0.23$ and $p=0.77$) nor an interaction between test condition (i.e., monaural versus bilateral) and ear of presentation ($p=0.39$ and $p=0.40$). Thus, the data were pooled across ear of presentation for these six listeners.

An additional analysis compared the magnitude of the perceptual-separation advantage across interferer types for the each listener group (and signal-processing condition for the NH listeners). Percentage-correct scores were converted to rationalized arcsine units (rau; Studebaker 1985) to offset the decrease in variance for very low (0-20%) or very high scores (80-100% correct). The perceptual-separation advantage was quantified for each interferer type and individual listener by calculating the performance difference between the monaural and bilateral conditions, averaged across the five TMRs common to all interferer types for each listener group. Separate repeated-measures analyses of variance – one for each listener group and signal-processing condition – were conducted on the estimated magnitude of the perceptual-separation advantage for the four interfering-talker conditions, with number and gender of the interfering talkers as within-subjects factors.

To keep with the standard convention in the literature, the perceptual-separation advantage was also calculated in dB. For each individual listener and masker type, the proportion correct for the monaural condition was derived from the psychometric-curve fit for a TMR of 0 dB, and the perceptual-separation advantage was quantified by calculating the TMR required to achieve the same level of performance in the bilateral condition. Estimates in dB were not performed for interferer types with non-monotonic performance functions in the monaural or bilateral condition. Furthermore, estimates in dB were not performed for the NH listeners presented with unprocessed speech where performance for a TMR of 0 dB was near ceiling level in many cases. The dB values were not analyzed statistically because they could not be calculated for all listeners in all conditions.

Follow-up experiment: out-of-set interferer keywords

One possible confounding factor with the experimental paradigm employed in the current study is that listeners might have been able to independently monitor the speech presented to the two ears to improve their performance in the speech-identification task. Because this was a closed-set task, and the interferers never spoke the same keywords at the target talker, it is theoretically possible that listeners could have identified the keywords spoken by the interferers in the second ear to rule them out as responses in the closed-set task. Thus, listeners could have achieved better performance in the bilateral condition than in the monaural condition even in the absence of interaural interactions.

A follow-up experiment examined the possible role that independent monitoring of the interferer signals presented to the second ear could have played in facilitating the performance improvements observed for the BI-CI and SSD-CI listeners. This experiment employed the same paradigm as in the main experiment, except that the interfering talkers produced keywords that were never spoken by the target talker and were not included as choices in the response set. This greatly reduced the possibility that independent monitoring of the contralateral interferers-only ear to narrow down the possible response choices could further improve performance. Five of the BI-CI listeners and three of the SSD-CI listeners from the main study participated. The experimental apparatus and procedure were the same as in the main experiment, except that fewer conditions were tested and a different set of interferer-signal recordings were used. Both groups were tested in the condition involving one same-gender interfering talker. The SSD-CI listeners were additionally tested with two same-gender interferers, because performance was close to ceiling in the single-interferer condition. The interferer stimuli were produced by the same set of talkers that produced the target-signal recording, but with color (“black” and “brown”) and number (“nine” and “ten”) keywords that were not included in the target response set.

RESULTS

Figure 1 plots the mean proportion of keywords correct (out of two for each trial, i.e., color and number) as a function of TMR for the three listener groups and, in the case of the NH listeners, for the three signal-processing conditions. For display purposes, the data were fit with a sigmoidal curve in most cases. In cases with non-monotonic performance functions, the data were instead fit with a third-order polynomial. Large stars denote conditions with a significant ($p < 0.05$) difference in performance between the monaural and bilateral conditions when the data were averaged across TMR. Asterisks denote individual TMRs where a significant ($p < 0.05$) difference in performance was observed. Figure 2 plots the mean magnitude of the improvement (in rau). Table 5 shows the magnitude of the improvement in dB, based on the horizontal distance between psychometric curves fit to the individual performance functions, similar to the fitted curves shown in Fig. 1 for the group-average data. (NH listeners in the unprocessed signal-processing condition were excluded from Table 5 because performance was near ceiling levels for a TMR of 0 dB, preventing a dB estimate of the improvement. Conditions with one interfering talker for the SSD-CI listeners and for the NH listeners in the SSD-vocoder condition were excluded from Table 5 due to non-monotonic performance functions.)

Presenting the interferers to both ears yielded a significant improvement in performance for at least a subset of interferer types for each of the three listener groups and for the three signal-processing conditions presented to the NH listeners. However, the pattern of results differed between the groups and signal-processing conditions in terms of which interferer types showed a significant increase in performance (large stars in Fig. 1), and in terms of the pattern of the effects of interferer number and gender on the magnitude of the performance increase (Fig. 2 and Table 5). The results for the NH listeners in the unprocessed condition are discussed first, followed by the results for the two CI listener groups, and finally the results for the two vocoder signal-processing conditions presented to the NH listeners.

NH listeners, unprocessed stimuli

The eight NH listeners presented with unprocessed speech (Fig. 1; top row) showed a significantly better performance in the bilateral condition than in the monaural condition for all five types of interferer ($p < 0.0005$). The magnitude of the improvement (Fig. 2A) was affected by both the number [$F(1,7) = 23.8$, $p < 0.005$] and the gender [$F(1,7) = 50.8$, $p < 0.0005$] of the interferers, with a larger improvement observed for two versus one, and for same-versus opposite-gender interferers. A significant interaction between number and gender [$F(1,7) = 8.89$, $p < 0.05$] reflected a larger effect of gender for the one-interferer conditions. The improvement in dB could not be precisely calculated for the NH listeners due to very high performance in the bilateral conditions; extrapolation of the measured performance functions yielded an estimated 15-20 dB of improvement for two same- or opposite-gender interferers.

BI-CI and SSD-CI listeners

The nine BI-CI listeners (Fig. 1, second row) showed a significant improvement for all four interfering-talker conditions ($p < 0.05$), but not for stationary noise ($p = 0.15$). The seven SSD-CI listeners (Fig. 1, third row) showed a significant improvement when the interferer(s) were of the same gender as the target talker ($p < 0.0005$ for one or two interferers), but not for opposite-gender interferers (one interferer: $p = 0.23$; two interferers: $p = 0.14$). There was also a small but significant ($p < 0.01$) reduction in performance when the interferer was presented to the CI ear in the stationary-noise condition (third row, right panel).

For the BI-CI listeners², the magnitude of the improvement was mainly determined by the number of interfering talkers [$F(1,7) = 1.04$, $p < 0.05$], with more improvement observed for one than for two interferers (Fig. 2B). The gender of the interferers did not affect the magnitude of the improvements: there was neither a significant main effect of gender ($p = 0.79$) nor an interaction between the gender and number of interferers ($p = 0.38$). For the SSD-CI listeners (Fig. 2C), the magnitude of the improvement was mainly determined by the gender of the interfering talkers [$F(1,6) = 19.2$, $p < 0.01$]. Although there was a trend toward more improvement for two same-gender interferers compared to one same-gender interferer (opposite the effect observed for the BI-CI listeners), neither the main effect of number ($p = 0.30$) nor the interaction between the gender and number of interferers ($p = 0.20$)

²One BI-CI listener (BI5) who experienced a significant decrease in performance in the bilateral conditions was excluded from the analysis of the magnitude of improvement across interferer types. See the discussion of intersubject variability at the end of this section describing the results for BI-CI and SSD-CI listeners.

was significant. For both CI listener groups, the maximum observed improvement across interferer types was 4-5 dB (Table 5) or 12-18 percentage points (Fig. 2).

Fig. 3 plots the magnitude of the perceptual-separation advantage for individual CI listeners. Each small circle and asterisk represents the magnitude of the perceptual source-separation advantage estimated for a given interferer type for one individual SSD-CI and BI-CI listener. (The open squares represent the mean data for the vocoder listeners, see below.) For both the BI-CI and SSD-CI listener groups, the magnitude of the improvement varied tremendously, with some listeners demonstrating a substantial perceptual source-separation advantage and some demonstrating little or no improvement. Binomial-logistic regression analyses for each listener and masker condition tested for the presence of interference rather than an advantage associated with the presentation of masker signals to the contralateral ear, with p-values adjusted for 60 multiple comparisons (16 listeners \times 5 masker types). One BI-CI listener (BI5) showed a large interference effect for all four interfering-talker conditions (Fig. 3, asterisks). This listener was pre- or peri-lingually deafened, but received her first implant at a relatively late age (51 yrs), and was clearly processing binaural information in a fundamentally different manner from the other BI-CI listeners (Nopp et al. 2004; Litovsky et al. 2010). Note that this listener was excluded from the calculation and analysis of the magnitude of the perceptual-separation advantage (Fig. 2A and Table 5; see Footnote 2).

NH listeners, vocoded stimuli

The magnitude of the perceptual source-separation advantage was greater for the vocoder conditions (Figs. 2D and 2E) than the average amount of advantage observed for the BI-CI and SSD-CI listeners (Figs. 2B and 2C). Expressed in dB, the vocoder listeners received, on average, a 7-15 dB of perceptual source-separation advantage which was substantially more than the average 4-5 dB demonstrated by the CI listeners (Table 5). However, when comparing the amount of perceptual source-separation advantage for individual CI participants (Fig. 3; small circles and asterisks) to the mean advantage for vocoded stimuli presented to NH listeners (Fig. 3; open squares) it was apparent that the best performing BI-CI and SSD-CI listeners experienced approximately the same amount of perceptual source-separation advantage as the average NH listener presented vocoded signals.

For the bilateral-vocoder conditions, the pattern of relative benefit between the interferer types was roughly similar to that for the actual BI-CI listeners, with a significant perceptual source-separation advantage observed for all four interfering-talker conditions (Fig. 1, fourth row; large stars). Figure 2D shows that there was a larger improvement in performance with one interfering talker than with two interfering talkers, as was observed for the BI-CI listeners (Fig. 2B). Although there was no main effect of the gender ($p=0.21$) or number of talkers ($p=0.09$) on the magnitude of the perceptual source-separation advantage, there was a significant interaction between gender and number of talkers [$F(1,7)=7.39$, $p<0.05$], supporting the observation that perceptual source-separation advantage was greater for one than for two same-gender interfering talkers. There was a very small (0.7 dB; Table 5), but significant improvement in performance in stationary-noise for the bilateral-vocoder conditions (Fig. 1; fourth row, right panel).

For the SSD-vocoder conditions, the pattern of results differed somewhat from the actual SSD-CI listeners, in that there was a significant perceptual source-separation advantage for all four interfering-talker conditions (large stars in Fig. 1, bottom row). Consequently, there was no significant main effect of gender ($p=0.13$) or interaction between number and gender ($p=0.17$) for the SSD vocoder conditions (Fig. 2E), in contrast to the significant main effect of gender observed for the SSD-CI listeners (Fig. 2C). There was, however, a significant main effect of number of interferers [$F(1,7)=10.8$, $p<0.05$], with a larger improvement observed for two than for one interferer (Fig. 2E), whereas there was only a non-significant trend in this direction for the actual SSD-CI listeners (Fig. 2C).

Follow-up experiment: out-of-set interferer keywords

The results of the follow-up experiment are plotted in Fig. 4 (circles and squares, with thick solid and dashed curves representing fits to the data). The fitted data from the main experiment with in-set maskers (for the subset of listeners that were also tested with out-of-set maskers) are replotted as thin curves in Fig. 4. Overall performance was better with out-of-set maskers (thick curves) than with in-set maskers (thin curves), reflecting the additional lexical cue available in the out-of-set case to identify which of the concurrent talkers was the target. Critically, both listener groups still showed a significant improvement in performance in the bilateral compared to the monaural condition ($p<0.05$) for all of the interferer types tested with out-of-set interferers, where no advantage from independently monitoring speech in the second ear would have been expected.

DISCUSSION

Overall, the results of this study suggest that a second auditory input can, in certain situations, improve the understanding of target speech in the presence of interfering talkers. NH listeners improved in all conditions, including the stationary-noise condition where listeners were expected to experience little difficulty perceptually separating the target and interfering sounds. This is consistent with previous results showing that these listeners can make use of detailed binaural information to improve speech understanding even with a simple stationary-noise interferer (Wan et al. 2010). NH listeners showed a larger improvement for two compared to one interferer, consistent with previous results showing that two interferers are more difficult to perceptually separate (Iyer et al. 2010). In addition, the NH listeners showed a larger improvement for same-gender compared to opposite-gender interferers, consistent with the idea that the relative lack of pitch and timbre differences with same-gender interferers would have made it more difficult to perceptually separate the target and interferers based on monaural cues (Brungart 2001).

For the BI-CI and SSD-CI listeners, the observed pattern of improvement suggests that the relatively crude signal provided by the CI in the second ear was sufficient to facilitate perceptual source separation in certain cases, although the improvements were much smaller compared to the NH listeners. For the BI-CI listeners, there was more improvement for one than for two interferers. For the SSD-CI listeners, there was more improvement for same-gender than for opposite-gender interferers. This contrast can be understood in terms of the different cues available in the target ear for each population. For the BI-CI listeners, the

target speech was presented to a CI ear. The CI does a poor job of relaying information about voice pitch, such that the perceived distinction between male and female voices is too subtle to facilitate their perceptual separation (Stickney et al. 2007). Thus, both same- and opposite-gender interferers are likely to produce substantial target-interferer similarity, making it difficult to perceptually separate the voices in the monaural condition. As a result, BI-CI listeners obtained the same improvement in performance in the bilateral condition, regardless of the gender of the interferers. The larger improvement for one compared to two interferers might reflect a reduction in the availability of dip-listening opportunities: as the number of interferers increases, there are fewer opportunities to hear the target speech during brief silences in the interfering speech waveform (Oxenham & Kreft 2014).

For the SSD-CI listeners, the target speech was presented to an acoustic ear. With acoustic hearing, the relative lack of pitch and timbre differences would have made it difficult to perceptually separate the target from same-gender interfering talkers in the monaural condition (Brungart 2001). As a result, the additional perceptual-separation cues available in the bilateral condition had the potential to improve performance. With opposite-gender interferers, SSD-CI listeners already had strong pitch and timbre cues to perceptually separate target and interferers in the monaural condition, such that the additional interaural cues available in the bilateral condition were less likely to produce a benefit. Thus for both CI listener groups, the second CI facilitated the perceptual separation when monaural cues were insufficient to do so.

It is not clear why there was no significant effect of the number of interferers on the magnitude of the advantage of SSD-CI listeners (although there was a non-significant trend toward more masking release for two than for one same-gender interferers). This result might reflect two offsetting effects. On one hand, listeners likely experienced a greater degree of difficulty in perceptually separating three (as compared to two) sources in the acoustic ear in the monaural condition, which would have increased the potential for a source-separation advantage in the bilateral condition, as was observed for the NH listeners (Fig. 1, top row). On the other hand, the two concurrent interferers are more likely to have been smeared by the CI processing in the second ear. This could have made it more difficult for the listener to fuse each interfering voice across the ears, thereby reducing the amount of perceptual-separation advantage in the bilateral condition.

Speech perception in competing-talker environments has been previously examined for BI-CI, but not for SSD-CI listeners. For BI-CI listeners, the current results contrast with two previous investigations that found little evidence of improved perceptual source separation with two CIs for spatially separated target speech and interferers (van Hoesel et al. 2008; Loizou et al. 2009). Our interpretation of these divergent results is that our study involved two factors that were critically important to observe a source-separation improvement. First, there was a dramatic difference between the signals presented to the two CIs, with the target stimulus completely absent from the second CI. Second, the speech corpus employed (the Coordinate Response Measure, or CRM) was specifically designed to make the perceptual separation of sound sources more difficult than for typical, everyday sentences because they employ an identical sentence structure (Brungart 2001). While van Hoesel et al. also employed the same dramatic difference between the signals presented to the two CIs, they

only used a noise interferer. Loizou et al. used estimates of interaural differences in the free field to generate relative target and masker levels in the two ears, yielding smaller differences between the ears than those in the current study. Furthermore, although they chose sentences from the same corpus (Institute of Electrical and Electronics Engineers sentences; Rothhauser et al. 1969) as both the targets and interferers, the target and interferers did not necessarily have identical sentence structure and were likely easier to perceptually separate than the CRM sentences. Thus, neither Loizou et al. nor van Hoesel et al. included both of these possibly critical factors in their study designs.

One more factor that could have contributed to the improved performance for BI-CI listeners in the current study is that performance was measured down to low (negative) TMRs (Freyman et al. 2008). Previous studies have used adaptive methods to measure the threshold TMR required for BI-CI listeners to achieve a given percentage-correct level of performance (e.g., van Hoesel et al. 2008; Loizou et al. 2009). This technique yielded threshold TMRs greater than 0 dB, where a level difference between target and masker signals can facilitate perceptual separation by providing a cue as to which of the competing signals is the target of interest (Brungart 2001). In the current study, the improvement for the BI-CI listeners was mainly observed for TMRs less than 0 dB (Fig. 1, middle row).

An issue that was not addressed in the current study but that might be important in producing a source-separation advantage for BI-CI or SSD-CI listeners is the potential role of CI experience or auditory training. In contrast to the results of van Hoesel et al. (2008) and Loizou et al. (2009), Eapen et al. (2009) found the presence of a binaural-interaction benefit in a group of bilateral CI listeners that were tested annually for four years after receiving their second CI. Like Loizou et al., Eapen et al. used simulations of the free-field listening to test spatial hearing. Unlike Loizou et al. who used speech maskers, Eapen et al. used noise maskers like van Hoesel et al. A binaural-interaction benefit was evidenced by improved performance when the worse ear – i.e., the CI on the side of the head closest to the noise – was added to the better ear. The fact that this benefit emerged after more than a year of listening experience raises the possibility that binaural advantages may emerge or increase with long-term BI-CI or SSD-CI listening experience or training (e.g., Tyler et al. 2010; Isaiah et al. 2014; Reeder et al. 2015). The current study demonstrated that a binaural-interaction benefit was much more easily observed for a situation involving the perceptual separation of interfering talkers than for a noise interferer. If BI-CI or SSD-CI listeners were provided with training in the perceptual separation of concurrent sources based on interaural-difference cues, an even larger advantage than that observed in the current study might emerge.

It should be noted that the paradigm employed in the current study is not a realistic representation of head shadow in the free field. In the bilateral condition, the ILD for the interferers was effectively 0 dB, but the target signal was completely absent from the second ear, yielding an effectively infinite ILD. In the free field, ILDs are never infinite. Head shadow can reach a maximum of about 20 dB for high-frequency signals in an anechoic environment (Feddersen et al. 1957; Bronkhorst & Plomp 1988). For broadband speech signals in the free field, maximum ILD values have been estimated to be as large as 5-8 dB for a far-field source in an anechoic environment (Culling et al. 2012; Kidd et al. 2005a), but

can be even larger for a near-field source at a distance less than 1 m (Brungart & Rabinowitz 1999). Using vocoder simulations in the same paradigm as that employed in the current study, Bernstein et al. (2015) examined the effect of non-infinite ILDs on contralateral unmasking by mixing an attenuated target signal in with the interferers presented to the contralateral ear. In the SSD-vocoder condition (simulating SSD-CI listening in NH listeners), a significant perceptual-separation advantage was observed for a target ILD as small as 6 dB. This result suggests that it is possible that similar benefits to those observed in the current study might also occur in the free field. Further studies will be required to determine whether the benefits observed here would translate to non-infinite ILDs and free-field conditions for BI-CI and SSD-CI listeners.

Although there was, on average, a significant perceptual-separation advantage observed in certain conditions, there was a large degree of intersubject variability apparent in the results (Fig. 4). This variability – consistent with the large degree of intersubject variability often observed in CI research – could be attributable to a number of possible demographic factors such as listening experience, duration of deafness, degree of residual hearing or limitations in the number of surviving spiral ganglion cells (Blamey et al. 2013; Long et al. 2014). This variability might also be attributable to particular distortions associated with CI processing, such as a lack of synchronization of the processors that can interact with the nonlinear compression algorithms to cause interaural disparities in stimulus level (van Hoesel 2012); distortions to perceived spatial location based upon processor mapping (Goupell et al. 2013; Fitzgerald et al. 2015); mismatch in the cochlear places of stimulation across the two ears due to differences in the surgical insertion depth (Kan et al. 2013, 2015; Goupell 2015); or limitations in spectral resolution due to current spread (Nelson et al. 2008). Even greater distortions of level and cochlear place of stimulation are likely for SSD-CI listeners than for BI-CI listeners because of the substantial differences in the processing of level (McDermott & Varsavsky 2009) and frequency (Stakhovskaya et al. 2007; Landsberger et al. 2015) in the CI and acoustic-hearing ears. For both listener groups, is it possible that under more favorable conditions – including greater neural survival, shorter duration of deafness, more listening experience, and CI processing that is optimized to facilitate binaural integration – an even larger perceptual source-separation advantage might be attained with CIs.

This view is supported by the results with vocoded stimuli presented to the NH listeners in experiment 1. Fig. 3 shows that the best performing BI-CI and SSD-CI listeners experienced approximately the same amount of perceptual source-separation advantage as the average listener presented vocoded signals. This suggests that the mean performance associated with vocoder simulations with eight frequency channels provide a reasonable estimate of the upper limit of the amount of perceptual source-separation advantage that might be experienced in a group of CI listeners, similar to the findings of previous investigations of speech perception in quiet and noise (Friesen et al. 2001). Examining the pattern of results across interferer types, the pattern of relative benefit across interferer types was roughly similar between the BI-CI listeners (Fig. 2B) and the bilateral-vocoder conditions (Fig. 2D). However, results for the SSD-CI listeners (Fig. 2C) and SSD-vocoder conditions (Fig. 2E) differed in that the vocoder results showed perceptual source-separation advantage even in the opposite-gender conditions. This suggests that under more favorable stimulus conditions and/or with increased physiological health of the implanted auditory pathway, it might be

possible for SSD-CI users to experience this advantage under conditions involving less similarity between the target and interferers. In fact, Fig. 3 shows that some individual SSD-CI listeners obtained a perceptual source-separation advantage with opposite-gender interferers, even though this was not the case in the group average (Fig. 1, third row).

It is an open question as to the nature of the interaural-difference cues that the CI listeners relied on to enhance the perceptual separation of target and masker signals. Presumably, the interferer signals were combined across the ears in some way to allow the listeners to better differentiate the target and interferers. One possibility is that, like the NH listeners, the BI-CI and SSD-CI listeners perceived the target and interferers as originating from different spatial locations. Neurons in the inferior colliculus are responsive to interaural differences presented to CIs for simple well-controlled stimuli (Hancock et al. 2012; Chung et al. 2014), facilitating sound localization and improving signal detection in noise for BI-CI listeners (Goupell & Litovsky 2015; Kan et al. 2015). However, with more complex signals like speech, signals presented with CIs may add interaural decorrelation, and thus tend to act like room reverberation, smearing localization cues and producing spatially diffuse images (Blauert & Lindemann 1986; Whitmer et al. 2012; Goupell et al. 2013; Goupell & Litovsky 2015). The paradigm employed in the current study had the potential to create dramatic differences in the perceived locations of the concurrent talkers, thereby overcoming the degraded binaural cues and blurry spatial images associated with CI processing to produce a binaural benefit. The monaural target is likely to have been perceived as originating at its ear of presentation. If listeners were able to perceptually fuse the interferers presented to the two ears, the interferers could have been perceived as a relatively punctate image originating from the center of the head, or a spatially diffuse image with a less-specific or non-specific spatial origin. In either case, a perceived spatial difference between the source locations could have given listeners a cue to distinguish the target and interferers (Freyman et al. 2001).

A second possibility, at least in the case of the SSD-CI listeners, is that the presentation of the interferers to the second ear could have resulted in a sound-quality difference between the target and interfering speech. If listeners fused the acoustic and electric versions of the interferers into a coherent sound image, this image is likely to have sounded like a mixture between natural and degraded speech, while the target signal is likely to have retained a purely acoustic quality. Listeners could have exploited such a difference in quality to perceptually separate the target and interferers. Consistent with this interpretation, Kidd et al. (2005b) investigated a scenario where the target and interfering speech were filtered into a series of non-overlapping narrow frequency bands. When a noise that was spectrally matched to the interferer was presented to the contralateral ear, performance improved. This result was interpreted in terms of a partial binaural fusion of the interfering speech in the target ear and the spectrally matched noise in the contralateral ear, giving the interferer a noise-like quality that allowed it to be more easily distinguished from the target speech.

A third possibility is that listeners might have listened separately to the signals presented to the two ears, developing a strategy to use information obtained about the interferers in the non-target ear to differentiate between the target and interferers presented to the target ear. This could have been done temporally, making use of the temporal cohesion between the

bilateral interferer signals to know when to listen for the monaural target. Temporal cueing could arise whether the interferer signals in the two ears were fused into a single perceived object or were processed separately. However, Bernstein et al. (2015) found that a one-channel vocoder, which preserves only gross envelope cues but not spectral information, did not yield an improvement in performance in the same paradigm employed here. Such a process could also have been carried out at a lexical level, by identifying the keywords spoken by the interferers in the second ear to rule them out as responses in the closed-set task. However, the follow-up experiment showed an improvement in performance for both listener groups in the bilateral condition even when the interfering talkers produced keywords that were not part of the response set (Fig. 4). This paradigm provided listeners with an additional cue to rule out response choices, but critically, the out-of-set cue was available in both the monaural and bilateral conditions. Because the out-of-set keywords could already be ruled out as possible response choices, independent monitoring of the speech information in the ear presented with only interferer(s) is unlikely to have provided any additional information to differentiate the target and interfering talkers. We cannot rule out the possibility that independent monitoring of the interfering voices presented to the second ear might have contributed to the benefit observed in the main experiment. In fact, in Fig. 4 the benefit appeared to be somewhat smaller with the out-of-set maskers (thick curves) than with in-set maskers (thin curves). This difference could be interpreted to imply a strategy of independent monitoring in the out-of-set condition. Alternatively, it might reflect a reduction in the amount of target-interferer confusion in the baseline monaural condition, thereby reducing the potential for improvement. In any case, the fact that a significant benefit was consistently observed in the bilateral condition even with out-of-set maskers suggests that the improvement in the main experiment was attributable, at least in part, to a source-separation advantage based on perceived spatial or quality differences between the target and interferers.

CONCLUSIONS

The results presented here demonstrate that for individuals with bilateral or single-sided deafness, providing bilateral input via cochlear implantation can facilitate the perceptual separation of concurrent sound sources. Previous evidence established that bilateral hearing via a CI improves speech perception in noise by allowing listeners more opportunities to take advantage of the ear with the better SNR. The current results show that for BI-CI and SSD-CI listeners, bilateral hearing via a CI can also restore the ability to make use of differences in the signals arriving at the two ears to more effectively organize an auditory scene.

ACKNOWLEDGEMENTS

Source of Funding:

This study was supported by a grant from the Defense Medical Research and Development Program (J.B.) and a grant from the NIH-NIDCD (R00 DC010206; M.G.). J.B. has a pending grant from the Med-El Corporation to support a clinical trial comparing cochlear implants to other existing treatment options for single-sided deafness.

J.B., M.J.G., and D.B. designed the experiments; J.B., M.J.G., G.S., and A.R. recruited participants and collected the data; J.B. analyzed the data; J.B., M.J.G., and D.B. wrote the paper. All authors discussed the results and implications and commented on the manuscript at all stages.

References

- Arbogast TL, Mason CR, Kidd G. The effect of spatial separation on informational and energetic masking of speech. *J Acoust Soc Am*. 2002; 112:2086–2098. [PubMed: 12430820]
- Arndt S, Aschendorff A, Laszig R, et al. Comparison of pseudo-binaural hearing to real binaural hearing rehabilitation after cochlear implantation in patients with unilateral deafness and tinnitus. *Otol Neurotol*. 2011; 32:39–47. [PubMed: 21068690]
- Bernstein JGW, Iyer N, Brungart DS. Release from informational masking in a monaural competing-speech task with vocoded copies of the maskers presented contralaterally. *J Acoust Soc Am*. 2015; 137:702–713. [PubMed: 25698005]
- Blamey P, Artieres F, Ba kent D, et al. Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiol Neurotol*. 2013; 18:36–47. [PubMed: 23095305]
- Blauert J, Lindemann W. Auditory spaciousness: Some further psychoacoustic analysis. *J Acoust Soc Am*. 1986; 80:533–542. [PubMed: 3745685]
- Bronkhorst AW, Plomp R. The effect of head-induced interaural time and level differences on speech intelligibility in noise. *J Acoust Soc Am*. 1988; 83:1508–1516. [PubMed: 3372866]
- Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am*. 2001; 109:1101–1109. [PubMed: 11303924]
- Brungart DS, Rabinowitz WM. Auditory localization of nearby sources. Head-related transfer functions. *J Acoust Soc Am*. 1999; 106:1465–1479. [PubMed: 10489704]
- Chatterjee M, Peng SC. Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hear Res*. 2008; 235:143–156. [PubMed: 18093766]
- Chung Y, Hancock KE, Nam SI, et al. Coding of electric pulse trains presented through cochlear implants in the auditory midbrain of awake rabbit: Comparison with anesthetized preparations. *J Neurosci*. 2014; 34:218–231. [PubMed: 24381283]
- Culling JF, Jelfs S, Talbert A, et al. The benefit of bilateral versus unilateral cochlear implantation to speech intelligibility in noise. *Ear Hear*. 2012; 33:673–682. [PubMed: 22717687]
- Darwin C, Brungart DS, Simpson BD. Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *J Acoust Soc Am*. 2003; 114:2913–2922. [PubMed: 14650025]
- Darwin C, Hukin R. Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *J Acoust Soc Am*. 2000; 107:970–977. [PubMed: 10687706]
- Dorman MF, Loiselle L, Stohl J, et al. Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear Hear*. 2014; 35:633–640. [PubMed: 25127322]
- Eapen RJ, Buss E, Adunka MC, et al. Hearing-in-noise benefits after bilateral simultaneous cochlear implantation continue to improve 4 years after implantation. *Otol Neurotol*. 2009; 30:153–159. [PubMed: 19180675]
- Erbele ID, Bernstein JGW, Schuchman GI, et al. An initial experience of cochlear implantation for patients with single-sided deafness after prior osseointegrated hearing device. *Otol Neurotol*. 2015; 36:e24–e29. [PubMed: 25406870]
- Fedderson WE, Sandel JT, Teas DC, et al. Localization of high-frequency tones. *J Acoust Soc Am*. 1957; 29:988–991.
- Firszt JB, Holden LK, Reeder RM, et al. Cochlear implantation in adults with asymmetric hearing loss. *Ear Hear*. 2012; 33:521–533. [PubMed: 22441359]
- Fitzgerald MB, Kan A, Goupell MJ. Bilateral loudness balancing and distorted spatial perception in recipients of bilateral cochlear implants. *Ear Hear*. 2015; 36:e225–e236. [PubMed: 25985017]
- Freyman RL, Balakrishnan U, Helfer KS. Spatial release from informational masking in speech recognition. *J Acoust Soc Am*. 2001; 109:2112–2122. [PubMed: 11386563]

- Freyman RL, Balakrishnan U, Helfer KS. Spatial release from masking with noise-vocoded speech. *J Acoust Soc Am*. 2008; 124:1627–1637. [PubMed: 19045654]
- Friesen L, Shannon R, Ba kent D, et al. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am*. 2001; 110:1150–1163. [PubMed: 11519582]
- Gallun FJ, Mason CR, Kidd G. Binaural release from informational masking in a speech identification task. *J Acoust Soc Am*. 2005; 118:1614–1625. [PubMed: 16240821]
- Gifford RH, Shallop JK, Peterson AM. Speech recognition materials and ceiling effects: Considerations for cochlear implant programs. *Audiol Neurotol*. 2008; 13:193–205.
- Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear Res*. 1990; 47:103–138. [PubMed: 2228789]
- Goupell MJ. Interaural envelope correlation change discrimination in bilateral cochlear implantees: Effects of mismatch, centering, and onset of deafness. *J Acoust Soc Am*. 2015; 137:1282–1297. [PubMed: 25786942]
- Goupell MJ, Kan A, Litovsky RY. Typical mapping procedures can produce non-centered auditory images in bilateral cochlear-implant users. *J Acoust Soc Am*. 2013; 133:EL101–EL107. [PubMed: 23363188]
- Goupell MJ, Litovsky RY. Sensitivity to interaural envelope correlation changes in bilateral cochlear-implant users. *J Acoust Soc Am*. 2015; 137:335–349. [PubMed: 25618064]
- Hancock KE, Chung Y, Delgutte B. Neural ITD coding with bilateral cochlear implants: Effect of binaurally coherent jitter. *J Neurophysiol*. 2012; 108:714–728. [PubMed: 22592306]
- Hansen MR, Gantz BJ, Dunn C. Outcomes after cochlear implantation for patients with single-sided deafness, including those with recalcitrant Meniere’s disease. *Otol Neurotol*. 2013; 34:1681–1687. [PubMed: 24232066]
- Hawley ML, Litovsky RY, Culling JF. The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *J Acoust Soc Am*. 2004; 115:833–843. [PubMed: 15000195]
- Hopkins K, Moore BCJ, Stone MA. Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *J Acoust Soc Am*. 2008; 123:1140–1153. [PubMed: 18247914]
- Isaiah A, Vongpaisal T, King AJ, et al. Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *J Neuroscience*. 2014; 34:11119–11130. [PubMed: 25122908]
- Iyer N, Brungart DS, Simpson BD. Effects of target-masker contextual similarity on the multitasker penalty in a three-talker diotic listening task. *J Acoust Soc Am*. 2010; 128:2998–3010. [PubMed: 21110595]
- Kan A, Litovsky RY, Goupell MJ. Effects of interaural pitch matching and auditory image centering on binaural sensitivity in cochlear implant users. *Ear Hear*. 2015; 36:e62–e68. [PubMed: 25565660]
- Kan A, Stoelb C, Litovsky RY, et al. Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear implant users. *J Acoust Soc Am*. 2013; 134:2923–2936. [PubMed: 24116428]
- Kidd G, Mason CR, Brughera A, et al. The role of reverberation in release from masking due to spatial separation of sources for speech identification. *Acta Acust Acust*. 2005a; 91:526–536.
- Kidd G, Mason CR, Gallun FJ. Combining energetic and informational masking for speech identification. *J Acoust Soc Am*. 2005b; 118:982–992. [PubMed: 16158654]
- Kidd, G.; Mason, CR.; Richards, VM., et al. Informational masking.. In: Yost, WA.; Popper, AN.; Fay, RR., editors. *Auditory Perception of Sound Sources*. Springer; New York, NY: 2008. p. 143-189.
- Landsberger DM, Svrakic M, Roland JT, et al. The relationship between insertion angles, default frequency allocations, and spiral ganglion place pitch in cochlear implants. *Ear Hear*. 2015 E-pub ahead of print.
- Litovsky RY, Goupell MJ, Godar S, et al. Studies on bilateral cochlear implants at the University of Wisconsin’s Binaural Hearing and Speech Laboratory. *J Am Acad Audiol*. 2012; 23:476–494. [PubMed: 22668767]
- Litovsky RY, Jones GL, Agrawal S, et al. Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans. *J Acoust Soc Am*. 2010; 127:400–414. [PubMed: 20058986]

- Litovsky RY, Parkinson A, Arcaroli J. Spatial hearing and speech intelligibility in bilateral cochlear implant users. *Ear Hear.* 2009; 30:419–431. [PubMed: 19455039]
- Loizou, PC. Speech processing in vocoder-centric cochlear implants.. In: Moller, A., editor. *Cochlear and Brainstem Implants.* Karger; Basel, Switzerland: 2006. p. 109-143.
- Loizou PC, Hu Y, Litovsky RY, et al. Speech recognition by bilateral cochlear implant users in a cocktail-party setting. *J Acoust Soc Am.* 2009; 125:372–383. [PubMed: 19173424]
- Long CJ, Holden TA, McClelland GH, et al. Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding. *J Assoc Res Otolaryngol.* 2014; 15:293–304. [PubMed: 24477546]
- Lu T, Litovsky R, Zeng FG. Binaural masking level differences in actual and simulated bilateral cochlear implant listeners. *J Acoust Soc Am.* 2010; 127:1479–1490. [PubMed: 20329848]
- Marrone N, Mason CR, Kidd G. Tuning in the spatial dimension: Evidence from a masked speech identification task. *J Acoust Soc Am.* 2008; 124:1146–1158. [PubMed: 18681603]
- Martin RL, McAnally KI, Bolia RS, et al. Spatial release from speech-on-speech masking in the median sagittal plane. *J Acoust Soc Am.* 2012; 131:378–385. [PubMed: 22280599]
- McDermott H, Varsavsky A. Better fitting of cochlear implants: Modeling loudness for acoustic and electric stimuli. *J Neural Eng.* 2009; 6:065007. [PubMed: 19850972]
- Nelson DA, Donaldson GS, Kreft H. Forward-masked spatial tuning curves in cochlear implant users. *J Acoust Soc Am.* 2008; 123:1522–1543. [PubMed: 18345841]
- Nopp P, Schleich P, D'haese P. Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants. *Ear Hear.* 2004; 25:205–214. [PubMed: 15179112]
- Oxenham AJ, Kreft HA. Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing. *Trends Hear.* 2014; 18:2331216514553783. [PubMed: 25315376]
- Peters R, Wyss J, Manrique M. Worldwide trends in bilateral cochlear implantation. *Laryngoscope.* 2010; 120:S17–S44. [PubMed: 20422715]
- Reeder RM, Firszt JB, Holden LK, et al. A longitudinal study in adults with sequential bilateral cochlear implants: time course for individual ear and bilateral performance. *J Speech Lang Hear Res.* 2014; 57:1108–1126. [PubMed: 24686778]
- Rothauer EH, Chapman WD, Guttman N, et al. IEEE recommended practice for speech quality measurements. *IEEE Trans Audio Electroacoust.* 1969; 17:225–246.
- Shannon RV, Zeng FG, Kamath V, et al. Speech recognition with primarily temporal cues. *Science.* 1995; 270:303–304. [PubMed: 7569981]
- Stakhovskaya O, Sridhar D, Bonham BH, et al. Frequency map for the human cochlear spiral ganglion: implications for cochlear implants. *J Assoc Res Otolaryngol.* 2007; 8:220–233. [PubMed: 17318276]
- Stickney GS, Assmann PF, Chang J, et al. Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences. *J Acoust Soc Am.* 2007; 122:1069–1078. [PubMed: 17672654]
- Studebaker GA. A “rationalized” arcsine transform. *J Speech Hear Res.* 1985; 28:455–462. [PubMed: 4046587]
- Tokita J, Dunn C, Hansen MR. Cochlear implants and single-sided deafness. *Curr Opin Otolaryngol Head Neck Surg.* 2014; 22:353–358. [PubMed: 25050566]
- Tyler RS, Witt SA, Dunn CC. Initial development of a spatially separated speech-in-noise and localization training program. *J Amer Acad Audiol.* 2010; 21:390–403. [PubMed: 20701836]
- van Hoesel RJM. Contrasting benefits from contralateral implants and hearing aids in cochlear implant users. *Hear Res.* 2012; 288:100–113. [PubMed: 22226928]
- van Hoesel R, Böhm M, Pesch J, et al. Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies. *J Acoust Soc Am.* 2008; 123:2249–2263. [PubMed: 18397030]
- Van Hoesel R, Litovsky RY. Statistical bias in the assessment of binaural benefit relative to the better ear. *J Acoust Soc Am.* 2011; 130:4082–4088. [PubMed: 22225062]

- Vermeire K, van de Heyning P. Binaural hearing after cochlear implantation in subjects with unilateral sensorineural deafness and tinnitus. *Audiol Neurotol.* 2009; 14:163–171.
- Wan R, Durlach NI, Colburn HS. Application of an extended equalization- cancellation model to speech intelligibility with spatially distributed maskers. *J Acoust Soc Am.* 2010; 128:3678–3690. [PubMed: 21218900]
- Whitmer W, Boyd A, Soraghan J, et al. Auditory externalization in hearing-impaired listeners: the effect of pinna cues and number of talkers. *J Acoust Soc Am.* 2012; 131:268–274.
- Zeitler DM, Dorman MF, Natale SJ, et al. Sound source localization and speech understanding in complex listening environments by single-sided deaf listeners after cochlear implantation. *Otol Neurotol.* 2015; 36:1467–1471. [PubMed: 26375967]
- Zurek, P. Binaural advantages and directional effects in speech intelligibility.. In: Studebaker, GA.; Hochberg, I., editors. *Acoustical Factors Affecting Hearing Aid Performance.* Allyn and Bacon; Boston, MA: 1993. p. 255-276.

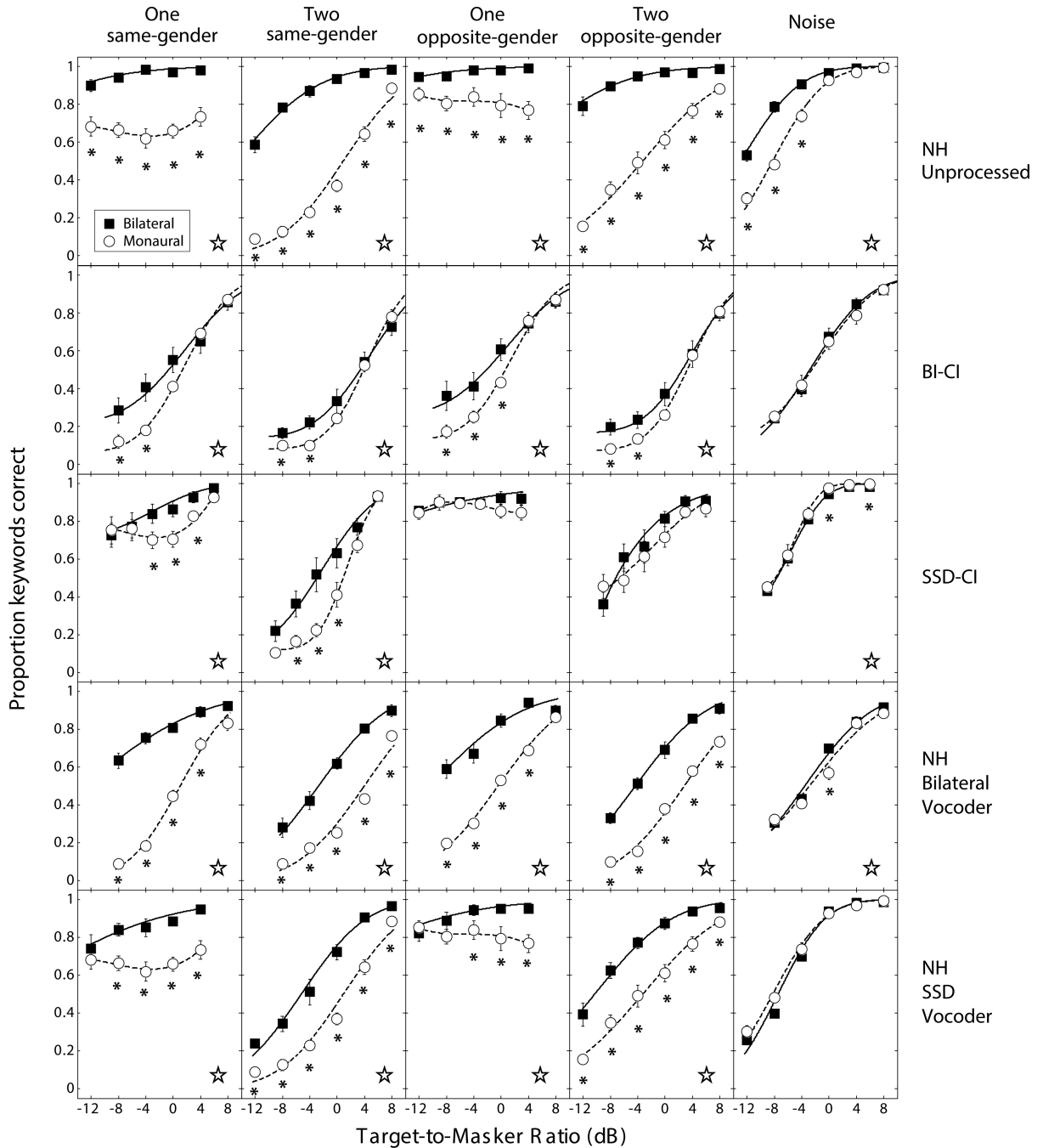


Fig. 1. Results of experiment 1 showing speech-identification performance as a function of TMR. The target signal was always presented monaurally. The interferers were presented monaurally to the same ear as the target (monaural condition) or diotically (bilateral condition). Conditions where bilateral performance was significantly better than monaural performance are identified by asterisks (for individual TMRs) and large open stars (for data pooled across TMRs). Error bars indicate ± 1 standard error of the mean across listeners.

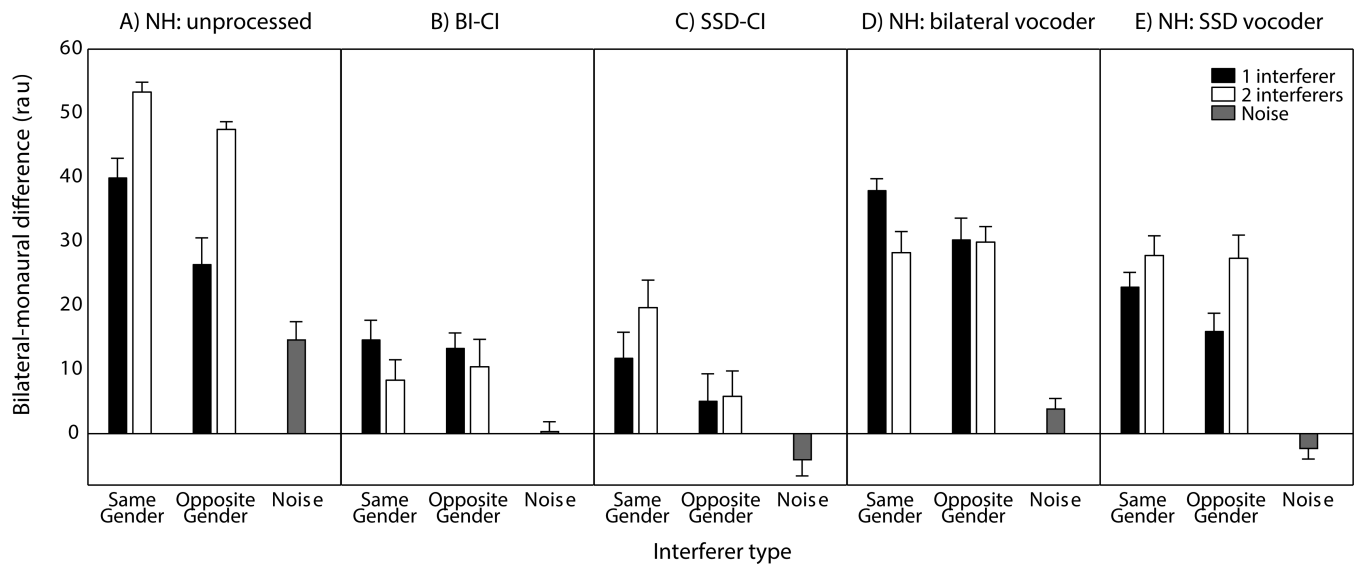


Fig. 2.

Results of experiment 1 showing the magnitude of the perceptual-separation advantage for each interferer type, averaged across TMR, for (a) the NH listeners presented with unprocessed stimuli, (b) the BI-CI listeners, (c) the SSD-CI listeners, (d) NH listeners in the bilateral-vocoded condition, and (e) NH listeners in the SSD-vocoded condition. Error bars indicate ± 1 standard error of the mean across listeners.

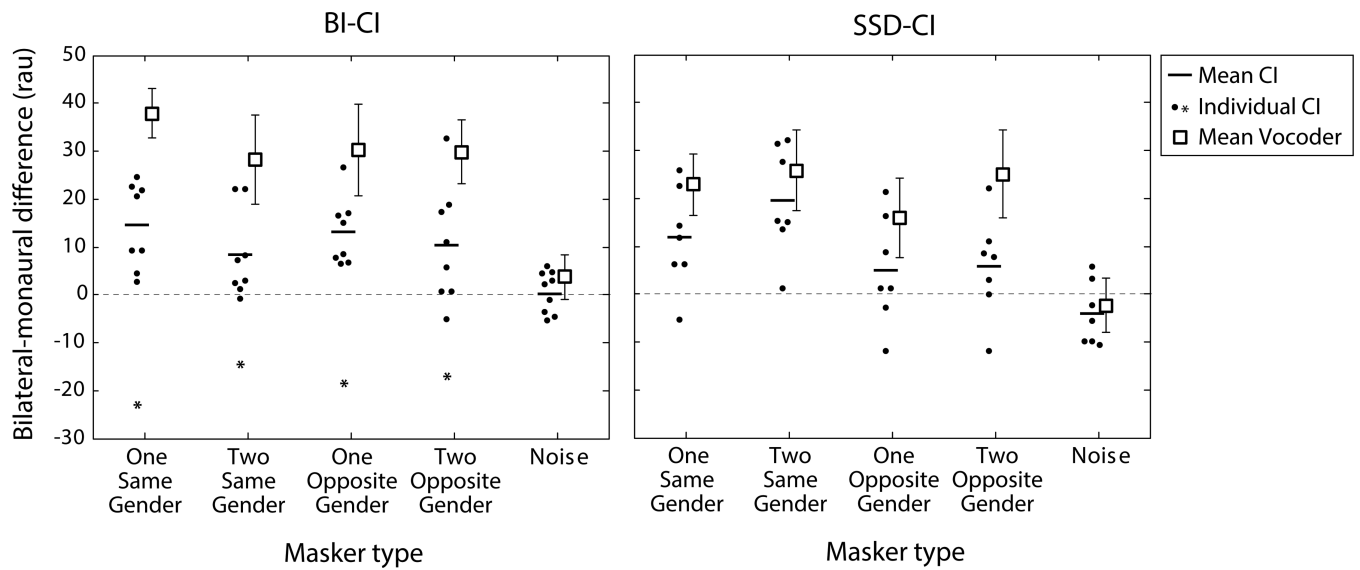


Fig. 3. Estimates of the magnitude of the perceptual-separation advantage in experiment 1 for individual CI listeners (small points) and mean data for the vocoder conditions presented to NH listeners (white squares). Asterisks indicate the BI-CI listener who demonstrated significant interference when the interferer signals were presented to the second CI; this listener was excluded from the computation of the group-mean average (horizontal bars). Error bars indicate ± 1 standard error of the mean across listeners.

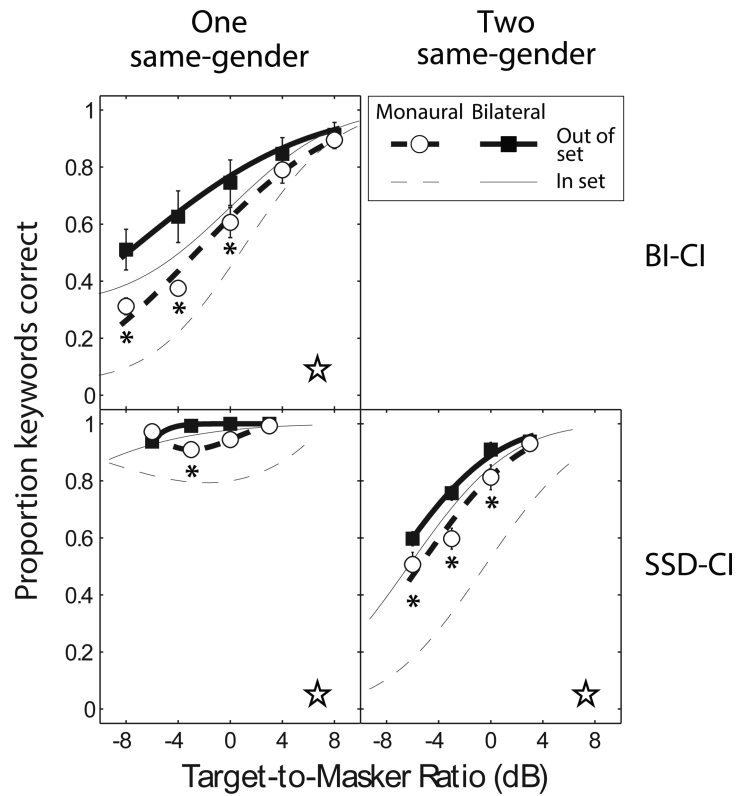


Fig. 4. Mean keyword-identification performance as a function of TMR for the subset of BI-CI and SSD-CI listeners tested in experiment 3. The color and number spoken by the interfering talkers (brown or black; nine or ten) were not items included in the response set, precluding the use of lexical information in the interferer-only ear to produce an improvement in performance relative to the monaural condition. Conditions where bilateral performance was significantly better than monaural performance are identified by asterisks (for individual TMRs) and large open stars (for data pooled across TMRs). Error bars indicate ± 1 standard error of the mean across listeners.

Table 1

Demographic information for the nine BI-CI participants.

| Listener | Sex | Age | Left CI Experience (years) | Right CI Experience (years) | Left CI Model | Right CI Model | Etiology |
|----------|-----|-----|----------------------------|-----------------------------|---------------|----------------|----------------------|
| BI1 | F | 61 | 5 | 7 | Freedom | Freedom | Hereditary |
| BI2 | M | 73 | 10 | 5 | Freedom | Nucleus 5 | Unknown |
| BI3 | F | 55 | 4 | 3 | Freedom | Freedom | Unknown |
| BI4 | M | 67 | 1 | 2 | Freedom | Nucleus 5 | Radiation |
| BI5 | F | 53 | 1 | 2 | Nucleus 5 | Nucleus 5 | Stickler's Syndrome |
| BI6 | F | 63 | 10 | 8 | Nucleus 5 | Nucleus 5 | Unknown |
| BI7 | F | 68 | 16 | 10 | Freedom | Freedom | Unknown |
| BI8 | F | 46 | 7 | 6 | Freedom | Nucleus 5 | Unknown |
| BI9 | F | 50 | 6 | 2 | Nucleus 5 | Nucleus 5 | Gradual hearing loss |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Demographic information for the seven SSD-CI participants.

| Listener | Sex | Age | CI Experience (months) | CI Model | CI Ear | Etiology | Acoustic Ear |
|----------|-----|-----|------------------------|----------------|--------|-------------------------------|--|
| SSD1 | M | 41 | 3 | Concert Flex28 | R | Enlarged Vestibular Aqueduct | Normal |
| SSD2 | M | 30 | 12 | Nucleus 422 | R | Sudden SNHL | Normal |
| SSD3 | M | 43 | 4 | Concert Flex28 | L | Sudden SNHL | Mild SNHL |
| SSD4 | M | 34 | 3 | Concert Flex28 | R | Sudden SNHL | Normal |
| SSD5 | F | 54 | 13 | Concert Flex28 | L | Cholesteotoma | Conductive loss |
| SSD6 | M | 52 | 19 | Nucleus 422 | L | Otosclerosis/ Surgical Trauma | Mild conductive loss & high-frequency SNHL |
| SSD7 | M | 46 | 6 | Concert Flex28 | L | Sudden SNHL | Mild SNHL |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Air-conduction audiometric thresholds (dB HL) for the seven SSD-CI listeners, and bone-conduction thresholds for the two SSD-CI listeners with conductive loss.

| Listener | Audiometric frequency (Hz) | | | | | | | |
|----------------------------|----------------------------|-----|------|------|------|------|------|------|
| | 250 | 500 | 1000 | 2000 | 3000 | 4000 | 6000 | 8000 |
| Air-conduction thresholds | | | | | | | | |
| SSD1 | 5 | 10 | 10 | 15 | 25 | 25 | 15 | 20 |
| SSD2 | 5 | 0 | 0 | -5 | 10 | 0 | 5 | 5 |
| SSD3 | 5 | 10 | 15 | 5 | 20 | 20 | 30 | 35 |
| SSD4 | 10 | 15 | 15 | 0 | 10 | 20 | 20 | 25 |
| SSD5 | 65 | 55 | 45 | 40 | 40 | 35 | 65 | 70 |
| SSD6 | 20 | 25 | 30 | 25 | 35 | 55 | 100 | 100 |
| SSD7 | 20 | 25 | 20 | 20 | 35 | 25 | 35 | 35 |
| Bone-conduction thresholds | | | | | | | | |
| SSD5 | 10 | 20 | 5 | 5 | | 15 | | |
| SSD6 | 0 | 10 | 10 | 10 | 35 | 40 | | |

Table 4

TMRs tested for each listener group and condition

| Listener Group | Processing condition | Interferer type | Min TMR (dB) | Max TMR (dB) | TMR step (dB) |
|----------------|----------------------|--------------------------|----------------|--------------|---------------|
| BI-CI | N/A | All | -8 | 8 | 4 |
| SSD-CI | N/A | One same-gender | -12 | 3 | 3 |
| | | All other interferers | -9 | 6 | 3 |
| NH | Unprocessed | One interferer | -12 | 4 | 4 |
| | | Two interferers or noise | -12 | 8 | 4 |
| | Bilateral vocoder | All | -8 | 8 | 4 |
| | | SSD vocoder | One interferer | -12 | 4 |
| | | Two interferers or noise | -12 | 8 | 4 |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Results from experiment 1 showing estimates of the magnitude of the perceptual source-separation advantage (dB) for CI listeners and for NH listeners presented with vocoded stimuli, derived from psychometric fits to the performance functions. The reported values were estimated at the performance level associated with a TMR of 0 dB in the monaural condition. Asterisks indicate conditions for which a dB estimate could not be computed due to a non-monotonic performance function.

| Group or Condition | Interferer type | One interferer | Two interferers |
|--------------------|-----------------|----------------|-----------------|
| BI-CI | Same-gender | 4.9 | 5.1 |
| | Opposite-gender | 5.1 | 3.2 |
| | Noise | 0.4 | |
| SSD-CI | Same-gender | * | 4.1 |
| | Opposite-gender | * | 2.2 |
| | Noise | -1.4 | |
| Bilateral vocoder | Same-gender | 15.3 | 7.1 |
| | Opposite-gender | 10.3 | 7.1 |
| | Noise | 0.7 | |
| SSD vocoder | Same-gender | * | 6.6 |
| | Opposite-gender | * | 7.0 |
| | Noise | -0.2 | |