Magnified interaural level differences enhance spatial release from masking in bilateral cochlear implant users

Richardson, Benjamin N.[1], Kainerstorfer, Jana M.[1,2],

Shinn-Cunningham, Barbara G.[1], and Brown, Christopher A.[3]

1. Neuroscience Institute, Carnegie Mellon University
2. Biomedical Engineering, Carnegie Mellon University
3. Department of Communication Science and Disorders, University of Pittsburgh

ABSTRACT

Bilateral cochlear implant (BiCI) usage makes binaural benefits a possibility for implant users. Yet, limited access to interaural time difference (ITD) cues and reduced saliency of interaural level difference (ILD) cues restricts perceptual benefits of spatially separating a target from masker sounds for BiCI users. Here, we explore whether magnifying ILD cues improves intelligibility of masked speech for BiCI listeners in a “symmetrical-masker” configuration, which controls for long-term positive target-to-masker ratio (TMR) at the ear nearer the target from naturally occurring ILD cues. We magnified ILDs by estimating moment-to-moment ITDs in 1-octave-wide frequency bands, and applying corresponding ILDs to the target-masker mixtures reaching the two ears at each time in each frequency band. We conducted two experiments, one with NH listeners using vocoded stimuli and one with BiCI users. ILD magnification significantly improved intelligibility in both experiments. BiCI listeners showed no benefit of spatial separation between target and maskers with natural ILDs, even for the largest target-masker separation. Because ILD magnification is applied to the mixed signals at each ear, the strategy does not alter the TMR in either ear at any time; improvements to masked speech intelligibility are thus likely from improved perceptual separation of the competing sources.

I. INTRODUCTION

Bilateral cochlear implant (BiCI) use has become more common in recent years, in part because of the potential of providing CI users access to binaural cues. This can help address poor spatial hearing outcomes in BiCI users, including less accurate performance than their normal-hearing peers on sound localization tasks (Dorman et al., 2016; Grantham et al., 2008; Jones et al., 2014) and on spatial release from masking (SRM) tasks (D’Onofrio et al., 2020; Loizou et al., 2009). These outcomes are due in part to poor sensitivity to interaural time differences (Noel and Eddington, 2013), which rely on a robust representation of temporal fine structure (TFS).

Historically, CIs have not provided a very robust representation of acoustic inputs (Lorenzi et al., 2006; Moore, 2008), especially interaural differences important in spatial hearing (Anderson et al., 2024; Grantham et al., 2008; Laback et al., 2015). Several CI stimulation strategies have attempted to re-introduce or extend ITD information in BiCIs (van Hoesel et al., 2009; Long et al., 2003; Srinivasan et al., 2020; Thakkar et al., 2023), or TFS more generally (Ausili et al., 2020). However, there are other limitations that reduce perceptual sensitivity, including current spread, frequency limitations on phase locking, and interaural asymmetry in electrode position (Anderson et al., 2024; Gray et al., 2021; Laback et al., 2015). In addition, reduced TFS sensitivity arising from the impaired auditory systems of CI users limits spatial hearing benefits even when devices are linked (Kan et al., 2015).

BiCI users do retain some sensitivity to interaural level differences (Grantham et al., 2008). In listeners with NH, ILDs can provide benefits to speech in a number of ways. When two competing sound sources arrive at a listener’s head from different directions, the target sound is closer to one ear than is the masker, resulting in a greater acoustic target-to-masker energy ratio (TMR) in that ear, termed the better-ear effect (Glyde et al., 2013a, 2013b). Even in situations where there is not a long-term average TMR benefit, ILDs can provide dynamic short term better-ear advantages (“glimpses”) at moments when one of the maskers has an energy “dip” (Bronkhorst and Plomp, 1992; Gibbs et al., 2022; Glyde et al., 2013a, 2013b). Beyond better-ear effects, ILD cues can also aid in the perception of spatial differences in competing sources, which can contribute to SRM by promoting sound segregation and supporting selective attention based on perceived source laterality (Best et al., 2005; Ihlefeld and Shinn-Cunningham, 2008; Middlebrooks and Waters, 2020). This suggests that naturally occurring ILDs can support SRM in NH listeners when target and masker sound sources are sufficiently spatially separated through both long-term and dynamic better-ear effects, and by promoting sound source segregation and enabling spatial selective attention.

BiCI configurations retain some ILD information, which means that listeners using such devices may benefit from a positive TMR at the ear nearer the target (Litovsky et al., 2009; Loizou et al., 2009). However, ILDs are reduced after CI processing (Dorman et al., 2014; Gray et al., 2021), likely due to a number of factors including dynamic range compression associated with frontend automatic gain control that operates independently at the two ears (Archer-Boyd and Carlyon, 2019, 2021; Spencer et al., 2019), backend mapping (Khing et al., 2013), and independent peak picking (Gray et al., 2021). With limited access to better-ear effects, bilateral CI users show both reduced glimpsing (Gibbs et al., 2022; Hu et al., 2018) and a reduced ability to selectively attend to a target in the presence of a spatially separated masker (Akbarzadeh et al., 2020; Goupell et al., 2016). As a result, natural ILD cues alone are not sufficient to support SRM in BiCI users (Ihlefeld and Litovsky, 2012).

Magnifying low-frequency ILD information (i.e., increasing ILD cue magnitudes to be larger than what occurs naturally) can significantly improve speech intelligibility in BiCI users compared to presenting naturally occurring, unprocessed ILDs (Brown, 2014). Specifically, for a target talker to the left of midline and a masker talker to the right, magnification of ILDs increases SRM (Brown, 2014). Because ILD magnification works on the target-masker mixture at each ear, it does not change TMR. This suggests that the benefits from ILD magnification may be due to improved spatial selection of the target stream. Although ILD magnification would require synchronization across devices, ILD cues are simple to introduce into existing BiCI strategies because they only require sensitivity to differences in stimulation intensity at each ear, which is better preserved in electric hearing than is timing information.

The current study extends these previous results by comparing different ILD magnification schemes, testing both speech intelligibility in NH listeners presented with vocoded stimuli (Experiment 1) and in BiCI users (Experiment 2). We used three ILD magnification conditions: No magnification, which used naturally occurring spatial cues via non-individualized head-related transfer functions (HRTFs; Gardner and Martin, 1994); Low-Frequency magnification, which applied the same HRTF-based natural cues, then applied ILD magnification in frequency bands below 2 kHz, similar to Brown (Brown, 2014); and Broadband magnification, which applied ILD magnification in frequency bands spread across a wider range of frequencies (125-5500 Hz). Crucially, the Broadband magnification approach extends magnified ILDs into mid-frequencies (i.e. 1-5 kHz), which are important for speech understanding (DePaolis et al., 1996).

We also changed the number of processing bands used. Though listeners benefit from magnification applied to 20 ERB-wide bands below 2kHz (Brown, 2014), we do not know how results would change if magnification was applied to fewer independent frequency bands. This is an important consideration if the strategy is to be implemented in devices, since fewer frequency bands reduces processor cycles, which in turn reduces processing latency and improves battery life. Future experiments will establish the fewest number of processing bands that robustly improves performance across various acoustic spatial scenes. The current study does not parametrically vary processing band number due to testing time constraints. However, as few as six one-octave processing bands significantly improved localization performance (Brown, 2018), far fewer bands than have been used previously. This led us to fix this number at four bands in the current study.

The current study also builds on previous iterations of ILD magnification by testing it in a symmetrical-masker configuration. We employed this configuration for a number of reasons. First, we wanted to observe the effectiveness of the strategy in a more acoustically complex scene than the single-masker configuration used previously, as a three-source environment is challenging both to the algorithm and to the listeners. We also wanted to control for better-ear acoustic benefits in asymmetrical target-masker configurations, which lead to improved long-term TMR (Glyde et al., 2013a) and allow shorter-term glimpsing (Bronkhorst and Plomp, 1992; Gibbs et al., 2022). In a symmetrical-masker configuration with the target at midline and a masker on either side, better-ear benefits do not increase as target-masker separation increases; instead, each masker moves closer to its ipsilateral ear while target position does not change. We do not believe that TMR improvements can explain any perceptual benefits from ILD magnification, because the algorithm is applied to the mixed signals at each ear, so that TMR does not change with magnification.

In NH listeners, we predicted that naturally occurring ILDs alone would be insufficient for SRM at small target-masker separations, but would be sufficient at large separations where ILDs are larger. We expected ILD magnification to improve SRM by increasing the perceived spatial separation of target and maskers, allowing more successful spatial selective attention. We expected Low-frequency magnification to yield better performance than naturally occurring ILDs and that Broadband magnification would produce the best performance (and the strongest SRM).

For BiCI listeners, we expected no SRM with naturally occurring ILDs (Brown, 2014; Ihlefeld and Litovsky, 2012), even for large target-masker separations. We hypothesized that Low-frequency ILD magnification would enhance spatial perception and support SRM, and that Broadband ILD magnification would produce even better perceptual spatial separation and greater SRM.

II. MATERIALS AND METHODS

A. Participants

Sixteen NH subjects (Experiment 1) and seven BiCI patients (Experiment 2) participated. A hearing screening with NH subjects confirmed that all participants had audiometric thresholds of 25 dB HL or lower at octave frequencies between 125 and

8000 Hz. BiCI users had at least 12 months of experience with both devices. During testing, they used their everyday program without modification, except for a minor adjustment to ensure that sounds were at a comfortable overall level. The ages and devices used by each BiCI user are shown in Table I. All participants were compensated for their participation either monetarily or with course credit. All procedures were approved by the Institutional Review Board of the University of Pittsburgh.

TABLE I. Age and device used in each ear, for each BiCI user.

Subject #	Age	Left Ear Device	Right Ear Device
1	68	Cochlear Nucleus 5	Cochlear Nucleus 5
2	72	Cochlear Nucleus 5	Cochlear Nucleus 5
3	76	Cochlear Nucleus 5	Cochlear Nucleus 5
4	75	MED-EL Opus 2	MED-EL Opus 2
5	64	Cochlear Nucleus 5	Cochlear Nucleus 5
6	52	AB Harmony	AB Harmony
7	72	Cochlear Nucleus 5	Cochlear Nucleus 5

B. Stimuli and Signal Processing

Target speech stimuli were drawn from the CUNY sentence corpus (Boothroyd et al., 1985) produced by a female talker. Masker stimuli were drawn from the IEEE corpus (“IEEE Recommended Practice for Speech Quality Measurements,” n.d.), which were produced by a different female talker. Masker sentences were concatenated if needed and then truncated to be equal in length to the target.

Stimulus processing and presentation, as well as experiment control, were accomplished using the Python programming language. All signal processing was performed offline (see schematization in Fig. 1). To generate stimuli, each of the two masker sentences were first adjusted in level relative to the target sentence level to achieve a TMR of +2 dB. Then, naturally occurring binaural cues were applied separately to the target and maskers by convolving each with appropriate non-individualized HRTFs (Gardner and Martin, 1994). Target sentences were simulated at midline, while the maskers were symmetrically positioned around midline at azimuths of ± 0, 15, 30, 45, 60, 75, or 90 degrees, all at ear height (0 degrees elevation). The 0-degree configuration (where target and maskers were all co-located) served as a reference. Although stimuli in the 0-degree ILD magnified conditions went through the same processing pipeline, neither ILD magnification scheme affected stimuli in the co-located, 0-degree configuration because estimated ITDs were at or near zero and thus, so were the applied ILDs (see below). Following HRTF spatialization, target and masker tokens were summed into left and right ear channels. Left and right ear channels were bandpass filtered between 125 and 5500 Hz to simulate the limited bandwidth of CI processing and to set the effective bandwidth of the signals to be comparable to that available to typically BiCI users (Exp 2).

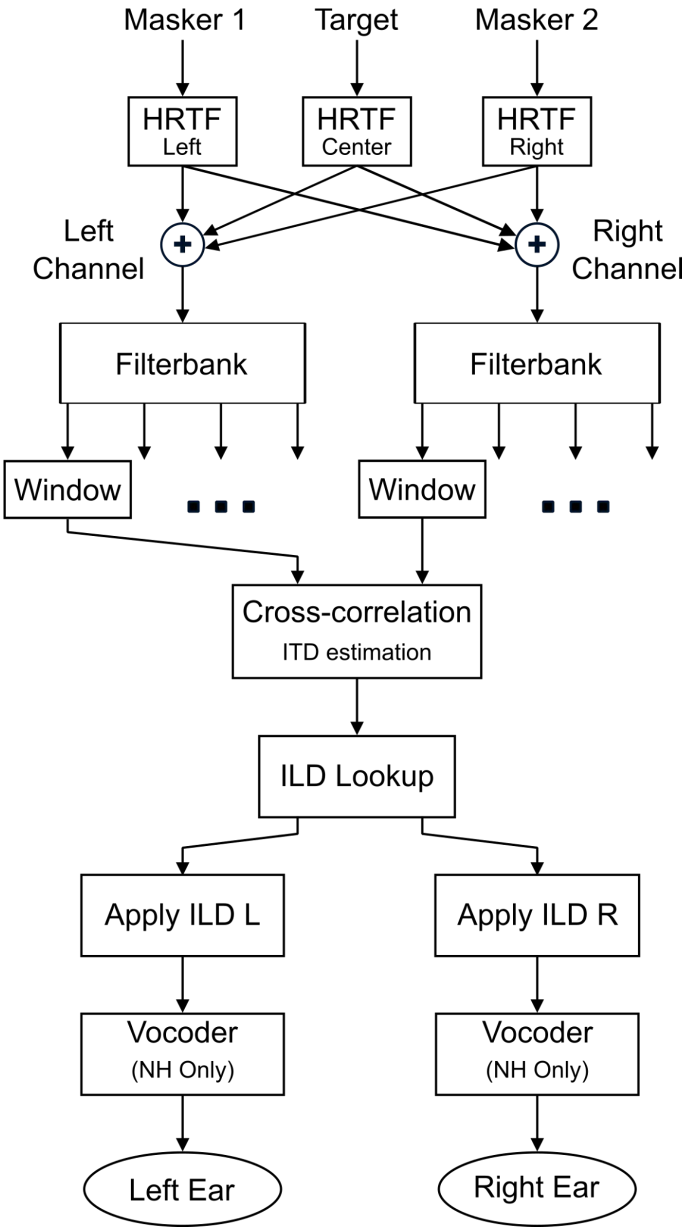

FIG. 1. Signal processing flow diagram. Target and masker sound stimuli were first spatialized using HRTFs. Then, working on the combined signal in the left and right ears, signals were filtered into frequency bands, and windowed into 20 ms bins. For each time-frequency bin, cross-correlation was used to estimate an ITD. This ITD was converted to an ILD with a lookup table, and applied via contralateral attenuation. Finally, the stimuli were vocoded (in Experiment 1 only).

In the No magnification condition, no further spatial processing was performed (stopped after “Bandpass Filter” in Fig. 1), and vocoding was applied (Exp 1 only). In the ILD magnification conditions, target-masker mixtures were filtered into different frequency bands, with the number depending on the processing condition (“Filterbank” in Fig. 1). We chose the cutoff frequencies to fix the number of frequency bands in which ILD magnification occurred to four. Both magnification filter banks used 4th-order Butterworth filters. In the Low-Frequency magnification condition, the filter bank comprised five contiguous bands with cutoff frequencies of 125, 250, 500, 1000, 2000, and 5500 Hz. ILD magnification was applied to the lowest four bands. In the Broadband magnification condition, four contiguous bands were created using cutoff frequencies of 125, 300, 900, 3500, and 5500, and ILD magnification was applied to all four bands, matching the number of processing bands in Low-freq magnification. As described in the Introduction, we were interested in whether fewer, wider bands would provide similar benefit to speech intelligibility as magnification using six independent frequency bands (Brown, 2018).

Following filtering, non-overlapping 20-ms boxcar-shaped windows of data were created within each frequency band (“Window” in Fig. 1). In each window, an ITD was estimated by finding the interaural delay corresponding to the lag of the maximum output of the cross-correlation of the left and right channels (“cross-correlation ITD estimation” in Fig. 1). A lookup table was used to convert the ITDs to ILDs. Specifically, a zero-μs ITD corresponded to a 0-dB ILD, a 750-μs ITD corresponded to a 32 dB ILD, and intermediate values were linearly interpolated, resulting in a linear weighting of approximately 0.043 dB/μs. The resulting ILD in each 20-ms segment was then achieved by attenuating the signal for the ear contralateral to the direction of the ITD for that frequency band and time window, leaving the ipsilateral ear segment unchanged. Although this processing approach causes intensity transitions from window to window that are heard as a light ‘static’ sound when listening to the broadband stimuli, pilot testing indicated that this static is not audible to either BiCI users or NH listeners presented with vocoded stimuli.

Finally, stimuli for NH listeners (Experiment 1) were processed with an 8-channel sinusoidal vocoder to make the available information similar to what a CI user could access. The first stage of the vocoder was a filterbank with cutoff frequencies logarithmically spaced between 125 and 5500 Hz. In each band, the amplitude envelope was extracted via half-wave rectification and low-pass filtering at the lesser of 400 Hz or half the bandwidth. These envelopes were then used to modulate zero-phase sine tone carriers with frequencies at the logarithmic center of the bands, which were then summed. Because the carriers were zero-phase, they delivered a constant ITD of 0 μs.

C. Equipment & Procedures

1. Experiment 1 (NH Listeners)

All digital signals (sampling frequency 44.1 kHz) were converted to analog signals with an RME Fireface UFX+ soundcard, and attenuated to an overall sound pressure level of 70 dB with Atlas AT100-RM passive attenuators. Listeners wore Etymotic ER-3a insert phones. All participants sat in an audiometric booth during testing while the experimenter sat in an acoustically isolated, but physically adjoined booth with a window between and an intercom allowing verbal communication.

Participants first heard ten target sentences presented in quiet to familiarize them with the normal acoustic voice of the target talker and then heard 10 vocoded target sentences. Participants were instructed to verbally repeat back as many of the words as possible produced by the centrally located “target” talker and to guess if they were not sure. Ten sentences, each with an average of five keywords (nouns, verbs, adjectives) for a total of 50 keywords, were presented in each ILD magnification (None, Low-freq, Broadband) and masker location (± 0, 15, 30, 45, 60, 75, and 90 degrees) condition. No participant ever heard a sentence more than once. For each condition, we quantified performance by calculating the total number of target keywords out of 50 that were correctly identified.

2. Experiment 2 (BiCI Users)

All equipment and procedures were identical to those used in Exp. 1, except for a few minor differences. Specifically, instead of a 70-dB SPL presentation level, BiCI participants used the attenuators to adjust the sound to a comfortable level. Instead of Etymotic ER-3 insert phones, direct-connect accessory cables connected to the CI speech processors delivered sound to participants. Participants used their everyday programs with no adjustments, including for automatic gain control. Finally, practice consisted of 10 unprocessed target sentences in quiet, with no vocoder practice.

III. RESULTS

A. Experiment 1

Fig. 2 presents the across-subject average results for Experiment 1, showing the raw percent correct speech intelligibility for each condition as a function of target-masker separation. In the co-located control condition, listeners correctly reported roughly 50% of the keywords. In the No magnification condition, performance was similar to the co-located configuration for less lateral maskers (±15, 30 & 45°), but was better for more lateral masker locations (±60, 75 & 90°). In contrast, for both of the ILD magnification conditions, performance was similar for all spatially separated masker conditions and higher than for the co-located configuration. Thus, target speech intelligibility is consistently better with ILD magnification than with naturally occurring ILDs when the maskers are spatially separated from the target but relatively close to midline.

A Shapiro-Wilk test of normality was conducted to determine whether percent correct data at spatially-separated masker locations were normally distributed. The results indicate that we fail to reject the null hypothesis ($p = 0.172$), supporting that the data is normally distributed. A 2-factor repeated-measures ANOVA was computed, with Masker location and ILD magnification as the two main factors and percent correct speech intelligibility as the dependent factor. Note that because the co-located conditions represented acoustically identical trials across ILD magnification conditions (i.e., all sources were at midline, so that ITD estimation and thus the applied ILD was always close to zero), data from this control condition were not included in the ANOVA analysis. Significant effects were observed for the main factors of ILD magnification

(F(2,30) = 5.4, p = 0.01, η^2 = 0.02) and Masker Location (F(5,75) = 3.9, p = .003, η^2 = 0.03); their interaction was also significant (F(10,150) = 2.1, p = 0.027, η^2 = 0.03). Post hoc two-tailed paired t tests with Bonferroni corrections revealed that performance in the Broadband condition was significantly better than in the No magnification condition (p=.018). However, performance in the Low-frequency magnification condition did not differ significantly from performance in either No magnification or Broadband magnification conditions ($p > .05$). To understand the significant interaction, we conducted follow-up one-way ANOVAs separately for each magnification condition. We found no significant effect of masker location for either the Low-Frequency or Broadband magnification conditions. However, there was a significant effect of spatial separation in the No magnification condition ($p < 0.001$). Pairwise t-tests showed that performance in the 15, 30, and 45 degree masker locations was significantly lower than in the 60 and 75 degree masker locations. This pattern of results clearly drives the interaction between target-masker separation and magnification conditions.

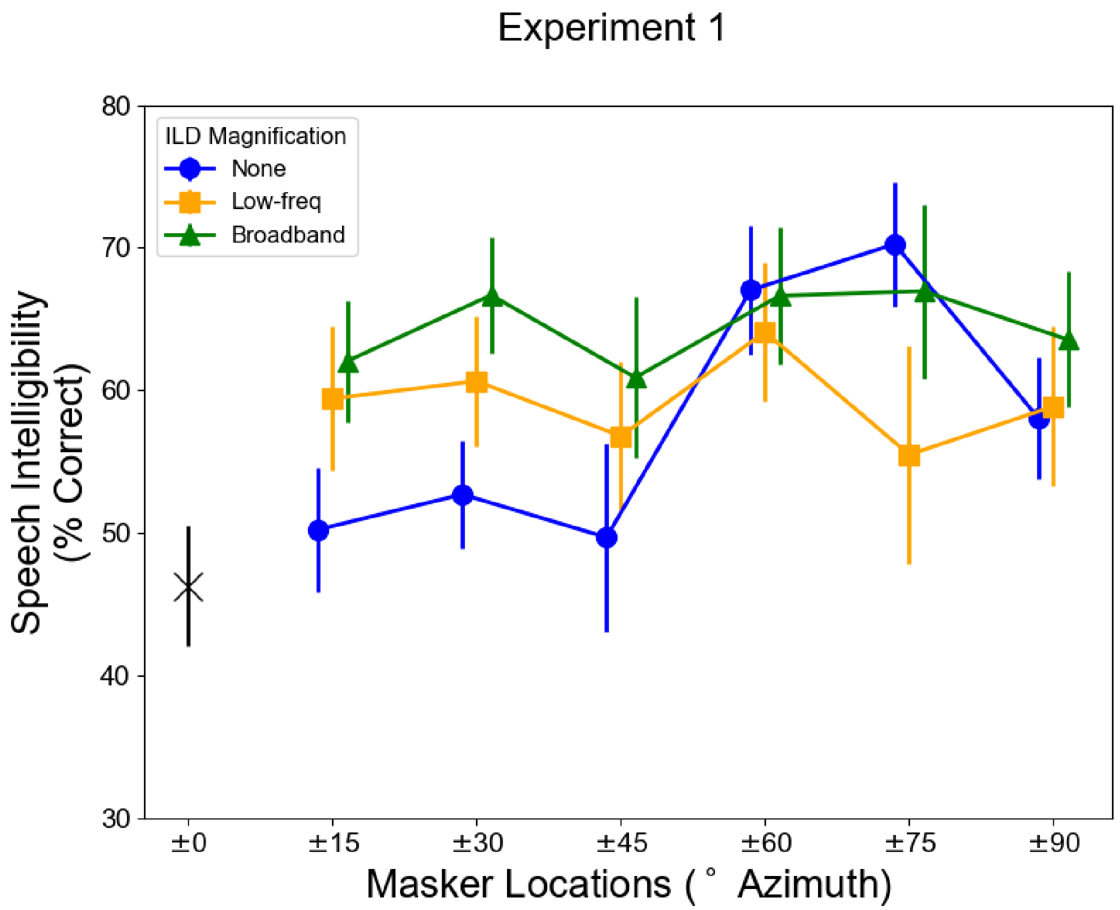

FIG. 2. Experiment 1 Results. Average percent keywords correct as a function of masker location for sixteen NH participants in a BiCI (vocoded) simulation. Data are shown for the three ILD magnification conditions: None (circles), Low-frequency (square), and Broadband (diamonds). The symbols show the across -subject average and the error bars show the across subject standard error.

B. Experiment 2

Fig. 3 presents the across-subject average results for Exp. 2 in 7 BiCI users, showing the raw percent correct speech intelligibility for each condition as a function of target-masker separation. In this population, performance in the co-located condition (about 40 percentage points) was generally lower than in the spatially separated conditions.

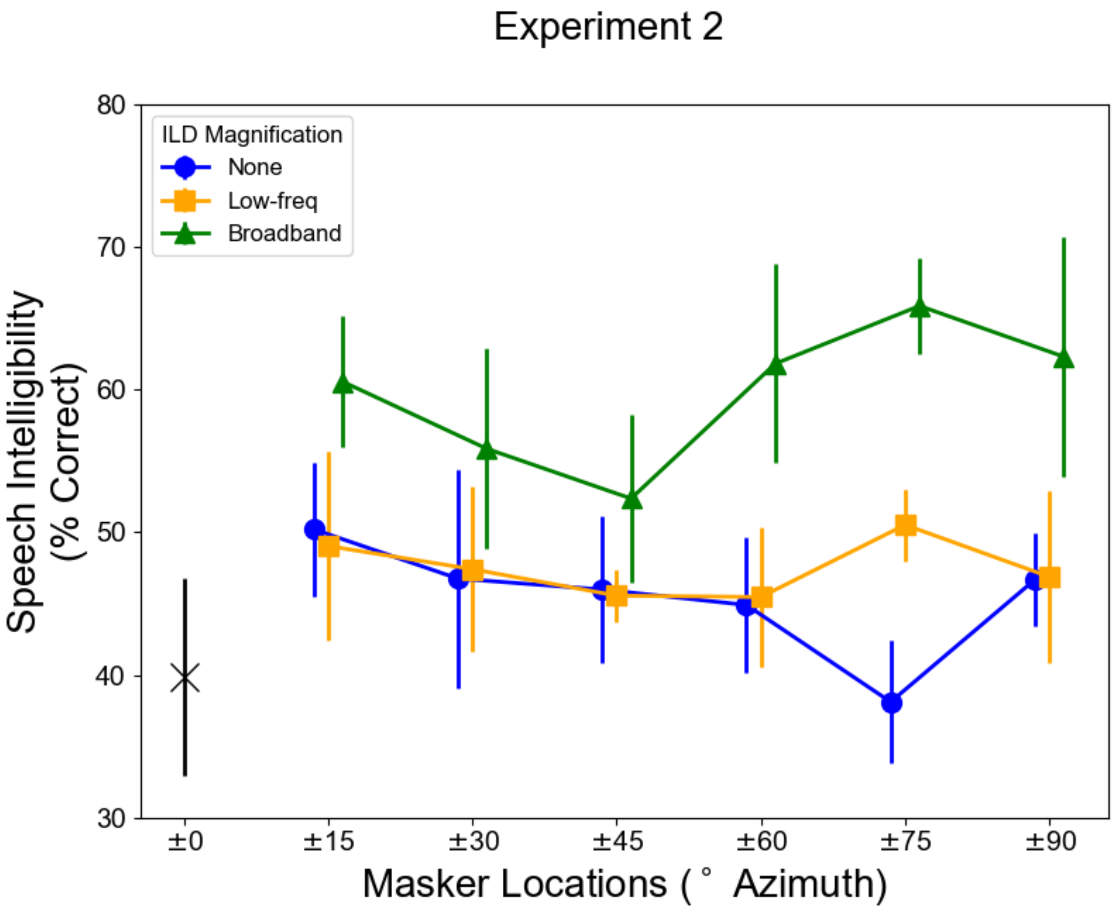

FIG. 3. Experiment 2 Results. Average percent keywords correct as a function of masker location for seven BiCI patients. Data are shown for the three ILD magnification conditions: None , low-frequency, and Broadband. The symbols show the across - subject average and the error bars show the across-subject standard error in each condition.

A Shapiro-Wilk test of normality of the performance data failed to reject the null hypothesis (p = 0.264), so we treated the data as normally distributed. A 2-factor repeated-measures ANOVA was conducted with dependent factors of masker location (for all spatially separated conditions) and ILD magnification, and percent-correct speech intelligibility as the dependent variable. The ANOVA revealed a significant main effect of ILD magnification ($F(2,12) = 11.4$, $p = 0.009$, $\eta^2 = 0.275$). However, neither Masker Location ($F(5,30) = 2.15$, $p = .116$, $\eta^2 = 0.174$) nor the interaction term ($F(10,60) = 1.18$, $p = 0.342$, $\eta^2 = 0.133$) were significant. Post hoc two-tailed paired t tests with Bonferroni correction for multiple comparisons revealed that the Broadband ILD magnification strategy produced performance that was significantly better than both No magnification ($p =0.041$), and Low-Frequency magnification ($p = .011$). However, there was no significant difference between performance in the No magnification and Low-Frequency magnification conditions ($p = 0.629$).

IV. DISCUSSION

Broadband ILD magnification significantly improved speech intelligibility in a symmetrical-masker task for both NH listeners (Experiment 1) and BiCI users (Experiment 2). Low-frequency magnification, however, did not have a significant effect on target sentence intelligibility. Given that our ILD manipulations should not appreciably change TMR or loudness of the stimuli, the benefit of ILD magnification may come from an increase in the perceived spatial separation of the target and maskers. To explore this possibility, we first analyzed the consequences of ILD magnification to rule out simple acoustic explanations for the effects we observed.

With naturally occurring ILDs, the better-ear effect contributes to improved speech intelligibility, and this effect grows as the separation between target and masker increases. One might reasonably wonder, then, whether introducing ILD magnification simply increases better-ear effects (i.e., increased TMR), rather than increasing perceived target-masker separation. However, as discussed earlier, the algorithm works by attenuating the summed target-masker mixture at each ear, as dictated by the ITD in that spectro-temporal bin. That is, both the target and masker in a given ear are attenuated by the same amount. Thus, the acoustic TMR at either ear is not altered, logically ruling out increased better-ear benefit.

We confirmed this by analyzing the processing, stimuli, and spatial configurations used in this study. We spatialized and summed stimuli as in the experiments, bandpassed and windowed the left and right channels, and calculated the estimated ITDs and associated ILDs (the “gain tracks”) to be applied at each time and in each frequency in each ILD magnification condition and for each masker azimuth. An example of these gain tracks is shown in Fig. 4. Panel A depicts the symmetrical-masker spatial configuration, with example waveforms of each source. Panel B plots the applied ILDs in each time-frequency bin for the Low-frequency (upper panel) and Broadband (lower panel) magnification conditions. The particular example shown placed maskers at ±60°. The waveforms in Panel A are time-aligned to the ILD tracks in Panel B. The legend at the bottom of the figure shows how ILD magnitudes correspond to the different colors in Panel B. The highest frequency band in the Low-frequency magnification plot (2 - 5.5 kHz) is solid blue, because ILD magnification was not applied to this band. It is visually apparent that when a given sound source is more intense than the others in the mixture, it drives the ITD estimate at that moment in time (for example, Masker 1 drives the ITD estimation at around 0.25 s, whereas the Target drives it at around 1.1 s).

Using the calculated gain tracks, we investigated whether TMR changed with the application of ILD magnification. In order to calculate TMR, we applied the ILD gain tracks to the unmixed target and masker signals independently (actual ILD magnification always works on the mixed signals). We then computed the TMR in each time-frequency window, and compared that TMR across the ILD magnification conditions. We found that the difference in TMR between naturally occurring ILD stimuli and ILD magnified stimuli at any given time-frequency bin for any stimulus was on the order of a rounding error (ie., less than about 1e-13). This was true for all masker azimuths: regardless of the magnitude of the level change in a given time-frequency bin, the target and masker were both equally attenuated in the contralateral ear, resulting in no change in TMR. Thus, changes in TMR cannot explain our results.

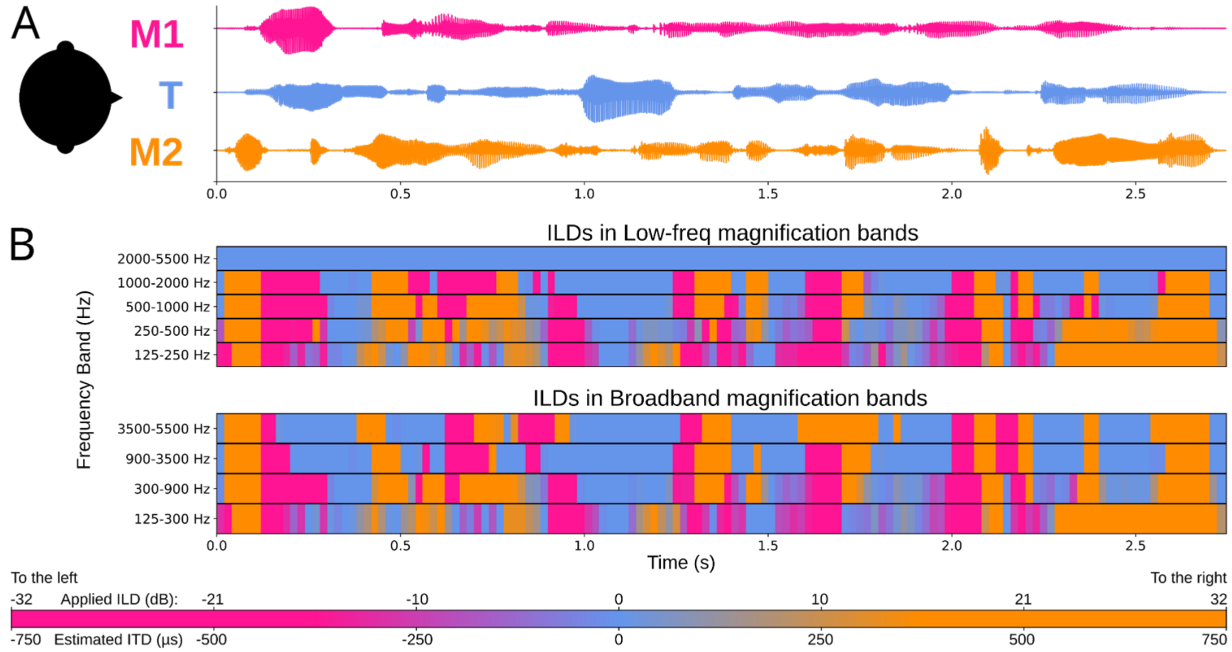

FIG. 4. Acoustic analysis example sentence. Panel A shows the relative positions of the target (T), masker 1 (M1) and masker 2 (M2), along with corresponding example waveforms. Panel B shows applied ILDs in spectro-temporal bins for Low-frequency (top plot) and Broadband (bottom plot) magnification. In this example, maskers were spatialized to +/-60 degrees. The legend at the bottom maps color to the applied ILD (and estimated ITD for reference). The waveforms in Panel A and the ILD plots in Panel B are time-aligned, making it easier to observe that when a given sound source is more intense than the others, that source biases the iTD estimation toward its azimuth. Eg., when M1 (spatialized to the left) is relatively high at around .2 s, the ILD applied is to the left. Note that the highest frequency band in Low-frequency magnification (2000-5500 Hz) always shows a 0 dB ILD, as no magnification was applied in this band.

Given that ILD magnification introduces changes in SPL, we also considered the possibility that ILD magnification resulted in a reduction in perceived loudness of the masker sounds, explaining the pattern of speech intelligibility results. However, because the sound source with the largest amplitude in a given mixture across the two ears drives the ITD estimate (see Fig. 4), the ear that is attenuated is nearly always the quieter ear. Studies of binaural loudness summation show that reductions in binaural loudness are considerably less than would be predicted from the acoustics (Epstein and Florentine, 2009). In the most extreme case, going from a diotic presentation to a monaural presentation, binaural loudness summation would result in a 3-dB change in energy; however, perceived changes in binaural loudness predict considerably smaller changes. These calculations suggest that changes in perceived loudness are too modest to explain the changes in performance we found with increased spatial separation.

Relatedly, it may be that the applied attenuation reduces the overall level of the maskers, reducing their ability to interfere with the target. The target would be less affected by the attenuation because it will always be unchanged in the unattenuated ear. This may be particularly true for BiCI users, for whom insertion depths, neural survival, and programming differences may lead to reduced binaural fusion. If a BiCI user is receiving two slightly different, unfused versions of the scene, then the presence of the masker in the to-be-attenuated ear may add interference, but not spatial separation because of the reduced fusion. In this case, the attenuation from the applied ILD may simply be reducing the intensity of what is essentially a monaural interferer. This possibility would not explain the performance of the NH group, however, who have good fusion with the approach we used. Given that the patterns of results across the two experiments track closely with one another, this possibility seems unlikely to be a significant factor in the outcome of the current study.

A. Experiment 1

We hypothesized that both Low-frequency and Broadband magnification would significantly improve performance compared to No magnification (ie., naturally occurring ILD cues). Supporting these hypotheses, we observed a significant benefit from Broadband magnification, and a small, albeit non-significant benefit from Low-frequency ILD magnification. These results are generally in line with a previous study that showed a significant increase in intelligibility even for low-frequency ILD magnification using different ILD magnification algorithm parameters and a different spatial configuration (Brown, 2014). The spatial configuration used here, with a masker to either side of the target, poses much greater challenges, both for the algorithm and for the listener, than when hearing a target to one side and a single masker to the other. Better-ear acoustic effects are limited in the symmetric masker configuration; moreover, ignoring two competing maskers is more difficult than ignoring a single masker. Together, these differences produce much more perceptual interference in the symmetrical masker configuration than the spatial configurations of this previous study, which may explain why there is no significant effect of Low-frequency ILD magnification. This more challenging listening situation seems to require greater spatial separation in higher frequency regions important for speech intelligibility (DePaolis et al., 1996; Hogan and Turner, 1998; Vickers et al., 2001). Consistent with this hypothesis, Broadband magnification yielded significantly greater benefit than Low-frequency magnification.

The only condition in which performance was not consistent across azimuth in this experiment was in the No magnification case. In this case, listeners showed a benefit at large target-masker separations, but at smaller target-masker separations, performance was essentially equivalent to that in the co-located configuration. This pattern of results suggests that in the challenging symmetrical masker configuration, small spatial separations are insufficient to allow spatial attention to work effectively. On the other hand, performance improved with Broadband ILD magnification for all spatially separated configurations, indicating that the magnification provides a perceptual benefit when natural cues are insufficient to support SRM.

B. Experiment 2

In BiCI users, we hypothesized that both Low-frequency and Broadband magnification would produce greater benefit from spatial separation between target and maskers than naturally occurring ILD cues. We observed a significant improvement in performance with Broadband magnification over No magnification; however, Low-frequency magnification provided no statistically significant benefit. The effect of ILD magnification was relatively consistent across azimuth in the BiCI group for all magnification conditions, as indicated by a lack of significant main effect of masker location and no interaction of masker location and magnification condition. In all three magnification conditions, percent correct did not vary significantly with spatial separation. It is worth noting that we observed significant effects of Broadband magnification here even though the subject pool comprised only 7 listeners. The lack of statistically significant differences between the No magnification and Low-frequency magnification may be due to a lack of power rather than a failure of ILD magnification.

To estimate how many subjects would be needed to observe an effect of Low-frequency magnification, we conducted a post-hoc power analysis using G*Power (Faul et al., 2007). This analysis suggested that 32 BiCI users would be needed to observe a difference between No magnification and Low-freq magnification with power of 0.80 and $\alpha = 0.05$.

Naturally occurring ILD cues provided insufficient perceptual separation to produce SRM, even with maskers at ±90°. Only in the magnified ILD conditions did spatial separation yield performance that was better than in the co-located configuration. This result supports our hypothesis that BiCI users require greater ILDs to perceive a separation between target and maskers. This is also consistent with the literature (Brown, 2014; Ihlefeld and Litovsky, 2012), and is likely due to the compressed perceptual space BiCI listeners experience for sources in the horizontal plane with naturally occurring cues (Grantham et al., 2007). ILD magnification can mitigate this limitation for BiCI users by expanding the perceptual space (Brown, 2018). While neither naturally occurring ILDs (No magnification) nor Low-frequency ILD magnification improved performance for our BiCI users compared to the co-located configurations, Broadband ILD magnification did. Given that our ILD magnification approach does not alter the TMR at either ear, these results strongly suggest that Broadband ILD magnification allows BiCI listeners to better focus on the target stream and suppress the maskers by deploying spatial selective attention. However, neither naturally occurring ILDs nor Low frequency ILD magnification allowed the BiCI users to focus spatial attention effectively.

C. General Discussion

Both NH and BiCI listeners benefitted from magnified ILD cues. For NH listeners, Broadband magnification improved performance compared to naturally occurring ILDs when symmetrically positioned maskers were close to midline and natural ILDs were too small to support SRM. For BiCI users, Broadband magnification significantly improved performance compared to both Low frequency magnification and No magnification conditions.

The overall benefit of ILD magnification observed in the current study is less than what was observed when listeners heard a target presented with a single masker (Brown, 2014). In this earlier non-symmetric masker study, ILD magnification increased percent correct performance by about 30 percentage points. In contrast, Broadband magnification in the current study improved performance by around 20 percentage points over no magnification. This difference likely reflects the perceptual difficulty associated with the symmetrical masker configuration used here. Another potential factor is the number of processing bands, which was reduced from 20 in the previous study to 4 here. Follow-up studies are needed to establish the relationship between processing band number and speech benefit. Nevertheless, 20 percentage points of masking release represents a substantial benefit and indicates that ILD magnification can be effective in relatively complex auditory scenes.

1. ILD Magnification Enhances Spatial Differences, Thereby Supporting Spatial Attention

The question remains as to the mechanism by which ILD magnification provides speech intelligibility benefit. We propose that magnified ILDs enhance perceived spatial separation between sound sources, allowing listeners to focus auditory spatial selective attention. NH listeners can use auditory spatial cues to selectively attend to a target amongst spatially separated maskers (Noyce et al., 2021; Shinn-Cunningham, 2017; Shinn-Cunningham and Best, 2008). This process occurs through coordinated activity in multiple areas of the brain, including prefrontal cortex, parietal cortex, and auditory cortex (Alho et al., 2014; Choi et al., 2014; Deng et al., 2019a, 2019b; Noyce et al., 2022). Spatial auditory attention is less effective in listeners with hearing loss; indeed, performance is inversely correlated with spatial discrimination thresholds (Bonacci et al., 2019; Dai et al., 2018). Similarly, limitations of electrical stimulation reduce CI users' ability to capitalize on naturally occurring spatial cues to direct selective attention (Akbarzadeh et al., 2020; Goupell et al., 2016). This underscores the need for processing algorithms like ILD magnification that facilitate greater perceptual segregation and more effective deployment of spatial attention.

2. Corrective ILD Magnification

The current approach maximizes the perceptual separation between target and masker, which provides significant SRM. But there are other spatial configurations in which the current strategy will likely be less effective, if not detrimental. Specifically, if a target and masker are spatially separated, but on the same side of midline, ILD magnification may decrease the perceptual separation between them. This is because there is a limit to the perceived lateral position of a sound source. Aggressive ILD magnification as was used in the current study may cause ipsilateral sources to be hyper-lateralized (sources perceived to be as far to the side as possible), which may actually lead to reduced perceptual separation between them.

Corrective ILD magnification (Brown, 2018) may represent a potential compromise. Whereas the ILD magnification strategy used in the current study was designed to maximize the perceptual separation between target and maskers for our configurations (with the target at midline and maskers to the sides), corrective processing is designed to minimize rms error between perceived and actual locations of the sources in the mixture. When tailored to individual BiCI patients, corrective ILD magnification significantly improves localization accuracy; two patients presented with this strategy exhibited localization performance on par with a group of NH listeners. Studies are planned to explore the balance that will likely need to be struck between maximizing perceptual benefit between target and masker as in the current study, and maximizing perceptual accuracy of source location (Brown, 2018).

3. Future Work

In addition to experiments designed to maximize benefit across different perceptual tasks, we also aim to explore the frequency-, azimuth-, and subject-specific benefits of magnified ILD cues in an SRM paradigm. This will include an examination of the effects of parameters such as the number of magnification bands, the cutoff frequencies of those bands, and the ITD-to-ILD mapping function (lookup table). Four magnification bands were used in this study to explore whether SRM benefits could be obtained using a small number of frequency bands, which is less computationally demanding than similar past approaches (Brown, 2014). Relatedly, the optimal processing bandwidth will also need to be established. Both the number of bands and the bandwidths that provide the maximum benefit may be frequency-, azimuth-, or subject specific.

The algorithm was specifically designed to manipulate ILDs without a priori knowledge of the location of sound sources, or which source is the target and which are maskers. It can be effective even in relatively complex acoustic environments like the symmetrical-masker configuration employed in the current study. It also works best with modulated maskers, which have proven to be more difficult for traditional noise reduction approaches. But there are spatial configurations that may pose a problem. For example, if a target and masker are spatially separated but on the same side of midline, ILD magnification may actually reduce the perceptual separation between them. Future work will explore this possibility.

We argue that the observed benefits for understanding target speech come from enhancing the perceived spatial separation between target and maskers. Future experiments can explicitly examine this hypothesis by measuring neural responses to sound in both normal hearing listeners and BiCI users completing spatial selective attention tasks with and without ILD magnification. If the magnified ILD cues used here do, in fact, allow for greater sound source segregation, this should be evident from neural signatures of spatial selective attention.

V. ACKNOWLEDGEMENTS

This work was supported in part by NIH grants R21-DC018408 to CAB and R01-DC015988 to BGSC. We would like to thank Sydney Sepkovic and Cece Restauri for help with data collection.

VI. REFERENCES

Akbarzadeh, S., Lee, S., and Tan, C.-T. (**2020**). “The Implication of Sound Level on Spatial Selective Auditory Attention for Cochlear Implant Users: Behavioral and Electrophysiological Measurement,” 2497–2501. Presented at the Proc. Interspeech 2020. doi:10.21437/Interspeech.2020-2836

Alho, K., Rinne, T., Herron, T. J., and Woods, D. L. (**2014**). “Stimulus-dependent activations and attention-related modulations in the auditory cortex: a meta-analysis of fMRI studies,” Hear. Res., **307**, 29–41. doi:10.1016/j.heares.2013.08.001

Anderson, S. R., Burg, E., Suveg, L., and Litovsky, R. Y. (**2024**). “Review of binaural processing with asymmetrical hearing outcomes in patients with bilateral cochlear implants,” Trends Hear., **28**, 23312165241229880. doi:10.1177/23312165241229880

Archer-Boyd, A. W., and Carlyon, R. P. (**2019**). “Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences,” J. Acoust. Soc. Am., **145**, 1389. doi:10.1121/1.5093623

Archer-Boyd, A. W., and Carlyon, R. P. (**2021**). “Further simulations of the effect of cochlear-implant pre-processing and head movement on interaural level differences,” J. Acoust. Soc. Am., **150**, 506. doi:10.1121/10.0005647

Ausili, S. A., Agterberg, M. J. H., Engel, A., Voelter, C., Thomas, J. P., Brill, S., Snik, A. F. M., et al. (**2020**). “Spatial hearing by bilateral cochlear implant users with temporal fine-structure processing,” Front. Neurol., **11**, 915. doi:10.3389/fneur.2020.00915

Best, V., Ozmeral, E., Gallun, F. J., Sen, K., and Shinn-Cunningham, B. G. (**2005**). “Spatial unmasking of birdsong in human listeners: Energetic and informational factorsa),” J. Acoust. Soc. Am., **118**, 3766–3773. doi:10.1121/1.2130949

Bonacci, L. M., Dai, L., and Shinn-Cunningham, B. G. (**2019**). “Weak neural signatures of spatial selective auditory attention in hearing-impaired listeners,” J. Acoust. Soc. Am., **146**, 2577. doi:10.1121/1.5129055

Boothroyd, A., Hanin, L., and Hnath, T. (**1985**). “A sentence test of speech perception: reliability, set equivalence, and short term learning,” Publications and Research, Retrieved from https://academicworks.cuny.edu/gc_pubs/399. Retrieved from https://academicworks.cuny.edu/gc_pubs/399

Bronkhorst, A. W., and Plomp, R. (**1992**). “Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing,” J. Acoust. Soc. Am., **92**, 3132–3139. doi:10.1121/1.404209

Brown, C. A. (**2014**). “Binaural Enhancement for Bilateral Cochlear Implant Users,” Ear Hear., **35**, 580–584. doi:10.1097/AUD.0000000000000044

Brown, C. A. (**2018**). “Corrective binaural processing for bilateral cochlear implant patients,” PLoS One, **13**, e0187965. doi:10.1371/journal.pone.0187965

Choi, I., Wang, L., Bharadwaj, H., and Shinn-Cunningham, B. (**2014**). “Individual differences in attentional modulation of cortical responses correlate with selective attention performance," Hear. Res., **314**, 10–19. doi:10.1016/j.heares.2014.04.008

Dai, L., Best, V., and Shinn-Cunningham, B. G. (**2018**). "Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention," Proceedings of the National Academy of Sciences, **115**, E3286–E3295. doi:10.1073/pnas.1721226115

Deng, Y., Choi, I., Shinn-Cunningham, B., and Baumgartner, R. (**2019a**). "Impoverished auditory cues limit engagement of brain networks controlling spatial selective attention," Neuroimage, **202**, 116151. doi:10.1016/j.neuroimage.2019.116151

Deng, Y., Reinhart, R. M. G., Choi, I., and Shinn-Cunningham, B. G. (**2019b**). "Causal links between parietal alpha activity and spatial auditory attention," (F. P. de Lange, H. Luo, and A. K. C. Lee, Eds.) Elife, **8**, e51184. doi:10.7554/eLife.51184

DePaolis, R. A., Janota, C. P., and Frank, T. (**1996**). "Frequency importance functions for words, sentences, and continuous discourse," J. Speech Hear. Res., **39**, 714–723. doi:10.1044/jshr.3904.714

D'Onofrio, 'onofrio Kristen, Richards, V., and Gifford, R. (**2020**). "Spatial Release From Informational and Energetic Masking in Bimodal and Bilateral Cochlear Implant Users," J. Speech Lang. Hear. Res., **63**, 3816–3833. doi:10.1044/2020_JSLHR-20-00044

Dorman, M. F., Loiselle, L. H., Cook, S. J., Yost, W. A., and Gifford, R. H. (**2016**). "Sound Source Localization by Normal-Hearing Listeners, Hearing-Impaired Listeners and Cochlear Implant Listeners," Audiology and Neurotology, **21**, 127–131. doi:10.1159/000444740

Dorman, M. F., Loiselle, L., Stohl, J., Yost, W. A., Spahr, A., Brown, C., and Cook, S.

(**2014**). “Interaural level differences and sound source localization for bilateral cochlear implant patients,” Ear Hear., **35**, 633–640. doi:10.1097/AUD.0000000000000057

Epstein, M., and Florentine, M. (**2009**). “Binaural loudness summation for speech and tones presented via earphones and loudspeakers,” Ear Hear., **30**, 234–237. doi:10.1097/aud.0b013e3181976993

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (**2007**). “G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences,” Behav. Res. Methods, **39**, 175–191. doi:10.3758/BF03193146

Gardner, B., and Martin, K. (**1994**). *HRTF Measurements of a KEMAR Dummy-Head Microphone* MIT Media Lab Perceptual Computing.

Gibbs, B. E., Bernstein, J. G. W., Brungart, D. S., and Goupell, M. J. (**2022**). “Effects of better-ear glimpsing, binaural unmasking, and spectral resolution on spatial release from masking in cochlear-implant users,” J. Acoust. Soc. Am., **152**, 1230. doi:10.1121/10.0013746

Glyde, H., Buchholz, J., Dillon, H., Best, V., Hickson, L., and Cameron, S. (**2013a**). “The effect of better-ear glimpsing on spatial release from masking,” J. Acoust. Soc. Am., **134**, 2937–2945. doi:10.1121/1.4817930

Glyde, H., Buchholz, J. M., Dillon, H., Cameron, S., and Hickson, L. (**2013b**). “The importance of interaural time differences and level differences in spatial release from masking,” J. Acoust. Soc. Am., **134**, EL147–EL152. doi:10.1121/1.4812441

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

doi:10.1121/1.4962378

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Haynes, D. S., and Labadie, R. F. (**2008**). “Interaural time and level difference thresholds for acoustically presented signals in post-lingually deafened adults fitted with bilateral cochlear implants using CIS+ processing,” Ear Hear., **29**, 33–44. doi:10.1097/AUD.0b013e31815d636f

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., and Haynes, D. S. (**2007**). “Horizontal-Plane Localization of Noise and Speech Signals by Postlingually Deafened Adults Fitted With Bilateral Cochlear Implants*,” Ear Hear., **28**, 524. doi:10.1097/AUD.0b013e31806dc21a

Gray, W. O., Mayo, P. G., Goupell, M. J., and Brown, A. D. (**2021**). “Transmission of Binaural Cues by Bilateral Cochlear Implants: Examining the Impacts of Bilaterally Independent Spectral Peak-Picking, Pulse Timing, and Compression,” Trends in Hearing, **25**, 23312165211030411. doi:10.1177/23312165211030411

van Hoesel, R. J. M., Jones, G. L., and Litovsky, R. Y. (**2009**). “Interaural Time-Delay Sensitivity in Bilateral Cochlear Implant Users: Effects of Pulse Rate, Modulation Rate, and Place of Stimulation,” J. Assoc. Res. Otolaryngol., **10**, 557–567. doi:10.1007/s10162-009-0175-x

Hogan, C. A., and Turner, C. W. (**1998**). “High-frequency audibility: benefits for hearing-impaired listeners,” J. Acoust. Soc. Am., **104**, 432–441. doi:10.1121/1.423247

Hu, H., Dietz, M., Williges, B., and Ewert, S. D. (**2018**). “Better-ear glimpsing with symmetrically-placed interferers in bilateral cochlear implant users,” J. Acoust. Soc. Am., **143**, 2128–2141. doi:10.1121/1.5030918

“IEEE Recommended Practice for Speech Quality Measurements,” (**n.d.**). , doi:

10.1109/IEEESTD.1969.7405210. doi:10.1109/IEEESTD.1969.7405210

Ihlefeld, A., and Litovsky, R. Y. (**2012**). "Interaural Level Differences Do Not Suffice for Restoring Spatial Release from Masking in Simulated Cochlear Implant Listening," PLoS One, **7**, e45296. doi:10.1371/journal.pone.0045296

Ihlefeld, A., and Shinn-Cunningham, B. (**2008**). "Disentangling the effects of spatial cues on selection and formation of auditory objectsa)," J. Acoust. Soc. Am., **124**, 2224–2235. doi:10.1121/1.2973185

Jones, H., Kan, A., and Litovsky, R. Y. (**2014**). "Comparing Sound Localization Deficits in Bilateral Cochlear-Implant Users and Vocoder Simulations With Normal-Hearing Listeners," Trends in Hearing, **18**, 2331216514554574. doi:10.1177/2331216514554574

Kan, A., Jones, H. G., and Litovsky, R. Y. (**2015**). "Effect of multi-electrode configuration on sensitivity to interaural timing differences in bilateral cochlear-implant users," J. Acoust. Soc. Am., **138**, 3826–3833. doi:10.1121/1.4937754

Khing, P. P., Swanson, B. A., and Ambikairajah, E. (**2013**). "The Effect of Automatic Gain Control Structure and Release Time on Cochlear Implant Speech Intelligibility," PLoS One, **8**, e82263. doi:10.1371/journal.pone.0082263

Laback, B., Egger, K., and Majdak, P. (**2015**). "Perception and coding of interaural time differences with bilateral cochlear implants," Hear. Res., Lasker Award, **322**, 138–150. doi:10.1016/j.heares.2014.10.004

Litovsky, R. Y., Parkinson, A., and Arcaroli, J. (**2009**). "Spatial hearing and speech intelligibility in bilateral cochlear implant users," Ear Hear., **30**, 419–431. doi:10.1097/AUD.0b013e3181a165be

Loizou, P. C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J., and Roland, P. (**2009**). "Speech recognition by bilateral cochlear implant users in a cocktail-party setting," J. Acoust. Soc. Am., **125**, 372–383. doi:10.1121/1.3036175

Long, C. J., Eddington, D. K., Colburn, H. S., and Rabinowitz, W. M. (**2003**). "Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user," J. Acoust. Soc. Am., **114**, 1565–1574. doi:10.1121/1.1603765

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (**2006**). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," Proceedings of the National Academy of Sciences, **103**, 18866–18869. doi:10.1073/pnas.0607364103

Middlebrooks, J. C., and Waters, M. F. (**2020**). "Spatial Mechanisms for Segregation of Competing Sounds, and a Breakdown in Spatial Hearing," Front. Neurosci., , doi: 10.3389/fnins.2020.571095. doi:10.3389/fnins.2020.571095

Moore, B. C. J. (**2008**). "The Role of Temporal Fine Structure Processing in Pitch Perception, Masking, and Speech Perception for Normal-Hearing and Hearing-Impaired People," J. Assoc. Res. Otolaryngol., **9**, 399–406. doi:10.1007/s10162-008-0143-x

Noel, V. A., and Eddington, D. K. (**2013**). "Sensitivity of bilateral cochlear implant users to fine-structure and envelope interaural time differencesa)," J. Acoust. Soc. Am., **133**, 2314–2328. doi:10.1121/1.4794372

Noyce, A. L., Kwasa, J. A. C., and Shinn-Cunningham, B. G. (**2021**). "Defining attention from an auditory perspective," WIREs Cognitive Science, **n/a**, e1610. doi:10.1002/wcs.1610

Noyce, A. L., Lefco, R. W., Brissenden, J. A., Tobyne, S. M., Shinn-Cunningham, B. G., and Somers, D. C. (**2022**). “Extended Frontal Networks for Visual and Auditory Working Memory,” Cereb. Cortex, **32**, 855–869. doi:10.1093/cercor/bhab249

Shinn-Cunningham, B. (**2017**). “Cortical and Sensory Causes of Individual Differences in Selective Attention Ability Among Listeners With Normal Hearing Thresholds,” J. Speech Lang. Hear. Res., **60**, 2976–2988. doi:10.1044/2017_JSLHR-H-17-0080

Shinn-Cunningham, B. G., and Best, V. (**2008**). “Selective attention in normal and impaired hearing,” Trends Amplif., **12**, 283–299. doi:10.1177/1084713808325306

Spencer, N. J., Tillery, K. H., and Brown, C. A. (**2019**). “The Effects of Dynamic-range Automatic Gain Control on Sentence Intelligibility With a Speech Masker in Simulated Cochlear Implant Listening,” Ear Hear., **40**, 710–724. doi:10.1097/AUD.0000000000000653

Srinivasan, S., Laback, B., Majdak, P., and Arnoldner, C. (**2020**). “Improving Interaural Time Difference Sensitivity Using Short Inter-pulse Intervals with Amplitude-Modulated Pulse Trains in Bilateral Cochlear Implants,” J. Assoc. Res. Otolaryngol., **21**, 105–120. doi:10.1007/s10162-020-00743-6

Thakkar, T., Kan, A., and Litovsky, R. Y. (**2023**). “Lateralization of interaural time differences with mixed rates of stimulation in bilateral cochlear implant listeners,” J. Acoust. Soc. Am., **153**, 1912–1923. doi:10.1121/10.0017603

Vickers, D. A., Moore, B. C., and Baer, T. (**2001**). “Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies,” J. Acoust. Soc. Am., **110**, 1164–1175. doi:10.1121/1.1381534