

# The Benefits of Increased Sensation Level and Bandwidth for Spatial Release From Masking

Kasey M. Jakien,<sup>1,2</sup> Sean D. Kampel,<sup>2</sup> Samuel Y. Gordon,<sup>2</sup> and Frederick J. Gallun<sup>1,2</sup>

**Objective:** Spatial release from masking (SRM) can increase speech intelligibility in complex listening environments. The goal of the present study was to document how speech-in-speech stimuli could be best processed to encourage optimum SRM for listeners who represent a range of ages and amounts of hearing loss. We examined the effects of equating stimulus audibility among listeners, presenting stimuli at uniform sensation levels (SLs), and filtering stimuli at two separate bandwidths.

**Design:** Seventy-one participants completed two speech intelligibility experiments (36 listeners in experiment 1; all 71 in experiment 2) in which a target phrase from the coordinate response measure (CRM) and two masking phrases from the CRM were presented simultaneously via earphones using a virtual spatial array, such that the target sentence was always at 0 degree azimuth angle and the maskers were either colocated or positioned at  $\pm 45$  degrees. Experiments 1 and 2 examined the impacts of SL, age, and hearing loss on SRM. Experiment 2 also assessed the effects of stimulus bandwidth on SRM.

**Results:** Overall, listeners' ability to achieve SRM improved with increased SL. Younger listeners with less hearing loss achieved more SRM than older or hearing-impaired listeners. It was hypothesized that SL and bandwidth would result in dissociable effects on SRM. However, acoustical analysis revealed that effective audible bandwidth, defined as the highest frequency at which the stimulus was audible at both ears, was the best predictor of performance. Thus, increasing SL seemed to improve SRM by increasing the effective bandwidth rather than increasing the level of already audible components.

**Conclusions:** Performance for all listeners, regardless of age or hearing loss, improved with an increase in overall SL and/or bandwidth, but the improvement was small relative to the benefits of spatial separation.

**Key words:** Aging, Bandwidth, Binaural hearing, Head-related-transfer-functions, Hearing aids, Hearing loss, Sensation level, Spatial release from masking.

(*Ear & Hearing* 2017;38:e13–e21)

## INTRODUCTION

Spatial release from masking (SRM) plays an important role in speech understanding. Normal-hearing (NH) adults perform better at speech recognition tasks involving multiple talkers when stimuli are spatially separated rather than colocated (Marrone et al. 2008a; Gallun et al. 2013; Glyde et al. 2013a). Because SRM has been found to be reduced for hearing-impaired (HI) individuals (Dubno et al. 2002; Best et al. 2012), researchers have attempted to improve audibility thereby increasing access

to important frequency-dependent cues, such as interaural time differences (ITDs) and interaural level differences (ILDs; Marrone et al. 2008a; Ahlstrom et al. 2009). Gelfand et al. (1988) equated stimulus intensity levels among listeners and found that individuals with presbycusis gained less benefit from binaural cues than age-matched controls. As there was no frequency-specific processing to ensure that subjects with hearing loss had access to high-frequency cues, the HI group's ability to make use of ILDs was potentially limited. Researchers have since aimed to equate audibility among listeners and enhance spatial cue perception by testing persons fitted with bilateral hearing aids. For example, Ahlstrom et al. (2009) found that older hearing-impaired individuals could gain up to 4dB of benefit in a spatial release task when unaided, and that the amount of benefit was reduced when only the low-frequency energy was amplified. Benefits increased slightly when all frequencies of the speech stimuli were amplified such that they were considered audible.

Glyde et al. (2013b) utilized virtual stimuli to examine the separate impact of ILDs and ITDs on SRM for a speech recognition task using symmetrically placed maskers. In that study, speech stimuli were processed so that three versions of a recognition task were created, each version providing different available interaural cues—one condition with both ITDs and ILDs, as well as one ILD-only condition and one ITD-only condition. While Glyde et al. found that either ITDs or ILDs provided sufficient information to attain SRM, results showed that more SRM could be achieved with only ILD cues compared to only ITD cues. These results imply that HI listeners may suffer decreased SRM due to limited access to high-frequency ILD cues.

There is evidence that is consistent with the hypothesis that variations in SRM may occur according to the bandwidth used to present speech stimuli, but most studies have found that low-pass filtering is less detrimental than high-pass filtering. Kidd et al. (2010) examined the effects of reduced bandwidth on SRM by filtering three CRM sentences (one target at 0 degree and two maskers either colocated or symmetrically positioned at  $\pm 15$  or  $\pm 90$  degrees) into four different frequency ranges (“broadband” filtered from 0 to 6 kHz, “low-pass” filtered at 1.5 kHz, “mid-frequency” band-pass filtered from 1.5 to 3 kHz, and “high-frequency” band-pass filtered from 3 to 6 kHz). Target-to-masker ratios (TMRs) at threshold were best for the broadband condition and worsened consistently going from the broadband to low-pass conditions, with the poorest performance occurring in the mid- and high-frequency conditions. Limiting the frequency region to below 1.5 kHz minimally reduced SRM, while larger reductions were found for the mid-to-high band-pass frequency conditions. It should be noted that all of the participants were young normal-hearing individuals, and thus it is possible that the relative importance of the various frequency regions is not consistent across age and hearing ability. To make specific recommendations for hearing aid design, it is essential to test the population for whom the devices are intended.

<sup>1</sup>Otolaryngology/Head & Neck Surgery, Oregon Health & Science University, Portland, Oregon, USA; and <sup>2</sup>Department of Veterans Affairs, Portland VA Medical Center, National Center for Rehabilitative Auditory Research, Portland, Oregon, USA.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially.

Two recent studies involving HI listeners (Moore et al. 2010; Levy et al. 2015) examined the effect of extended bandwidth on speech recognition performance in a competing speech task with both collocated and spatially separated stimuli processed through filters with a range of upper cutoff frequencies. Both studies found that increasing the cutoff frequency in the region between 4 and 7.5 kHz provided the most benefit in spatially separated conditions. Ahlstrom et al. (2013) found that SRM and consonant recognition scores in speech-shaped noise improved for listeners wearing hearing aids when the cutoff frequency of the stimuli was increased from 3.4 to 7.1 kHz. While these studies examined the effects of extending speech stimuli cutoff frequencies beyond 4 kHz, the present study aimed to examine the effects of reduced bandwidth (low-pass filtering at 2 kHz) on speech intelligibility and SRM. These manipulations are not an attempt to explore the benefits of extending hearing aid bandwidths, but rather are potentially relevant to the benefits of meeting current amplification targets at standard frequencies.

One motivation for examining access to high-frequency information even in the range normally available in current hearing aid technology is that the speech intelligibility index (SII; ANSI 3.5, 2007) weighs lower frequencies more heavily than higher frequencies for speech intelligibility. Ahlstrom et al. (2013) found that an SII-like analysis accurately predicted thresholds for HI listeners in some conditions. This is consistent with the Kidd et al. (2010) finding that removing low-frequency information was more detrimental than was removing high-frequency information. The present study is an attempt to more directly relate the findings of Kidd et al. (2010) to the observed benefits of extending bandwidth.

With regard to the general importance of amplification across frequency, Gallun et al. (2013) found an improvement in listener performance (i.e., average thresholds and overall best performance achieved) when an equal sensation level (SL) target was used rather than a fixed SPL. Similarly, Arbogast et al. (2005) found that SRM for a highly informational masker improved for both NH and HI individuals when higher masker SLs were used rather than lower masker SLs. For the NH group in that study, mean masker SLs were 38.8 dB (high SL) or 25.8 dB (low SL); SRM improved from 12.3 dB on average (low SL) to 18.7 dB on average (high SL). In a speech intelligibility task using two or four maskers, Best et al. (2013) found that HI listeners experienced reduced SRM compared with NH listeners. The researchers hypothesized that this was due to a limit on release from energetic masking for those with impaired hearing, as was also hypothesized by Marrone et al. (2008c). However, as the researchers used linear hearing aid amplification (NAL-RP) which does not restore full audibility, it would have been the case that stimuli were presented at lower SLs for the HI group. Consequently, it is possible that the lower SL, which also would have increased energetic masking (and perhaps reduced speech comprehension) may have contributed to the reduced SRM.

The previous research has focused primarily on the benefit of extending bandwidth, but it is also important to understand the benefit of increasing the SL of those components that are made audible. The present study aimed to look more closely at the relationship between SRM, bandwidth, and the presented SL of target and maskers. We include in our participant group listeners of a wider age range and hearing status than have typically been tested to ensure that the data are generalizable to the larger population of hearing aid users.

## MATERIALS AND METHODS

### Experiment 1

**Participants** • Thirty-six individuals participated in experiment 1 (age range 18 to 77 years, mean 38.67 years.). All participants had bilateral pure-tone averages (PTAs; 0.5, 1, 2, and 4 kHz) below 37 dB hearing level (HL; mean 12.1 dB  $\pm$  8.2 dB HL) and all had fairly symmetrical hearing at 2 kHz and below (most had differences between the ears of less than 10 dB at all frequencies, and none had differences exceeding 20 dB). Twenty-nine of these subjects were members of the group of 52 participants described in Gallun et al. (2013). Note that the hearing loss and asymmetries present in this sample are greater than those in the sample reported in Gallun et al. due to the addition of 7 participants with greater hearing loss. Average audiograms across all 36 participants are shown in Figure 1A, and the relationship between age and PTA is shown in Figure 1B. For comparison, the same data are plotted for the participants in experiment 2.

Participants were tested in a sound-attenuated booth at the National Center for Rehabilitative Auditory Research (NCRAR) in Portland, OR. Completion of testing took between 1 and 4 two-hr visits per participant. All participants were monetarily compensated for their time. All procedures were approved by the VA Portland Health Care System Institutional Review Board.

**Stimuli** • Stimuli were chosen from the standard coordinate response measure corpus (CRM; Bolia et al. 2000), which is low-pass filtered at 8 kHz. CRM sentences take the form of “Ready (CALLSIGN), go to (COLOR) (NUMBER) now” and consist of eight possible call signs: (Arrow, Baron, Charlie, Eagle, Hopper, Laker, Ringo, Tiger) and 12 keywords: four colors (red, green, white, and blue) as well as the numbers 1 to 8. Four female and 4 male talkers speak all possible combinations of the call signs, colors, and numbers, supplying 256 total possible combinations of call signs and key words.

A virtual spatial array was created using the methods of Gallun et al. (2013) to simulate collocated and spatially separated

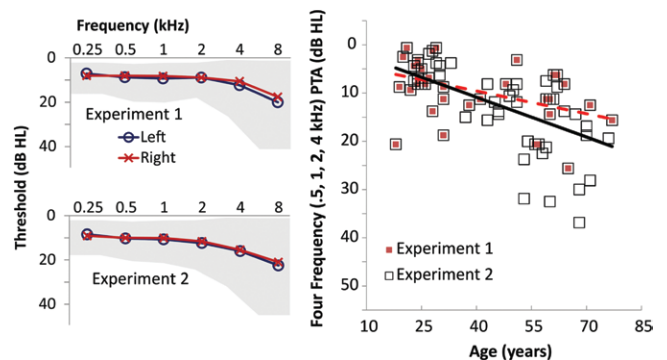


Fig. 1. Average audiometric thresholds for participants in experiments 1 and 2, shown as audiograms in (A) and as binaural PTAs based on four frequencies (0.5, 1, 2, and 4 kHz) in (B). The left and right ear average thresholds are plotted separately in (A) and are averaged together in (B). B. Also includes best-fitting lines for the 36 participants in experiment 1 (filled squares; dashed line) and the 71 participants in experiment 2 (open squares; solid line). Note that because all participants in experiment 1 were also included in experiment 2, (B) depicts those 36 as large square symbols with a smaller filled center square and the remaining 35 as large open squares. PTA indicates pure-tone average.

target and masking speech conditions via earphones and present stimuli with appropriate ITDs and ILDs for each spatial condition (Culling et al. 2004; Xie 2013; Glyde et al. 2013b). The approach used here was to use head-related impulse response measurements measured for a binaural manikin available from the CIPIC HRTF database (<http://interface.cipic.ucdavis.edu/sound/hrtf.html>) downloaded from the Music and Audio Research Laboratory, at New York University ([http://marl.smusic.nyu.edu/wordpress/projects/hrir\\_repository](http://marl.smusic.nyu.edu/wordpress/projects/hrir_repository)). Custom MATLAB functions were used to convolve each CRM wavefile to be presented with left ear and right ear head-related impulse response measurements. The source distance was 1 m.

**Procedures** • The adaptive tracking method described in Gallun et al. (2013) was used to determine listener thresholds for each testing condition in experiment 1. Participants listened to stimuli over Etymotic ER2 insert earphones and were asked to identify the target CRM phrase (the color-number combination after the callsign “Charlie”, always at 0 degree azimuth angle) in the presence of CRM maskers (either colocated or at  $\pm 45$  degrees). Responses were obtained using a computer monitor located in front of the participant. Participants initiated each track, selected their answers with a computer mouse, and were given feedback (“correct” or “incorrect”) after each answer. Performance thresholds were documented as TMRs. SRM was determined as the difference between thresholds in the colocated condition (target and maskers at 0 degree) and in the spatially separated condition. Although stimuli for experiment 1 were originally presented with both male and female target and masking voices, as described in Gallun et al. (2013), the current analysis focuses only on the stimuli consisting of male target and male maskers. This is to avoid potential confounds of audibility of the binaural cues associated with the target and maskers due to the differences in the frequency content of male and female voices. Performance was evaluated for the two spatial

conditions with the target sentence fixed at either 19.5 dB SL (“low SL” condition) or 39.5 dB SL (“high SL” condition). The levels of the two maskers were adjusted relative to level of the target sentence and were changed in level by 5 dB using a one-up/one-down adaptive tracking algorithm (Levitt 1971). Thresholds were determined as the average TMR across eight reversals. The level of the maskers was not allowed to exceed 85 dB SPL; however, this did not impact the adaptive tracks for any participants.

**High SL condition:** For the high SL condition, the target was presented using earphones at 39.5 dB SL, defined relative to each listener’s in-quiet speech reception threshold (SRT). SRTs were determined by the standard clinical test in which speech levels are reduced until a listener can correctly report 50% of the words spoken. SRT values were transformed from HL to dB SPL by adding 12.5 dB (ANSI S3.6 2004), then adding 39.5 dB to the value to determine the fixed level of the target sentence for each listener. Presenting the stimuli at 39.5 dB SL had improved TMRs, yet it prevented the participation of listeners with hearing loss great enough that this stimulus level would exceed their comfortable loudness threshold. Many of the high SL data were presented in Gallun et al. (2013) yet they have been updated with additional listeners and included in this paper as a means of comparing data with results from the low SL condition.

**Low SL condition:** Stimuli were presented with the target sentence at 19.5 dB SL above each listener’s SRT. For the comparisons presented here, only listeners who participated in both the high and low SL conditions are included.

## RESULTS

Figure 2 shows the TMR thresholds for both the low SL conditions (A and B) and high SL conditions (C and D). Figure 2A, C

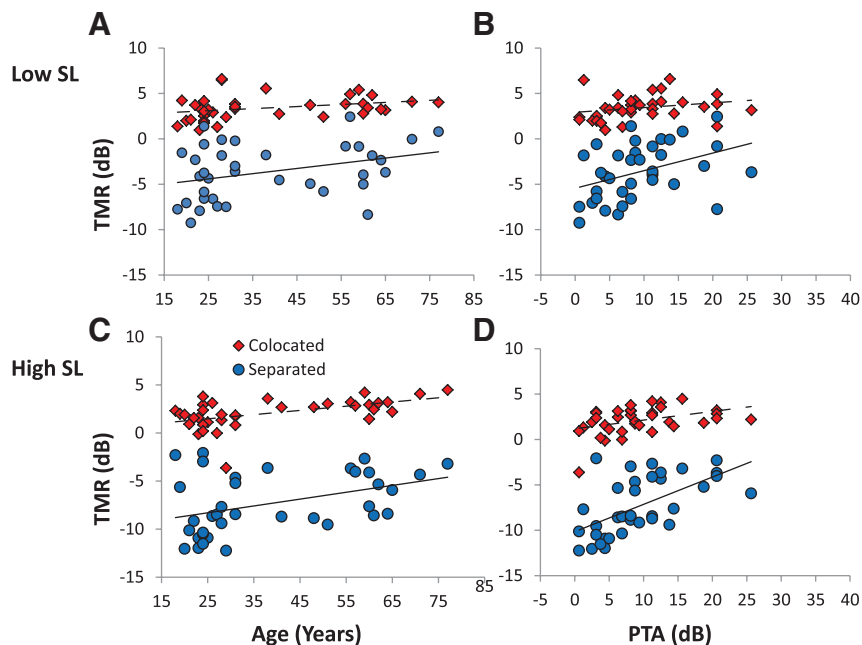


Fig. 2. Individual thresholds for the participants in experiment 1 are plotted for the low SL condition in (A) and (B) and the high SL condition in (C) and (D). Dashed lines and diamonds indicate performance when target and maskers were colocated, while solid lines and circles indicate performance under conditions of spatial separation. Correlations with age are shown in (A) and (C) and with binaural four-frequency (0.5, 1, 2, and 4 kHz) PTA in (B) and (D). PTA indicates pure-tone average; SL, sensation level.



**TABLE 1. Mean thresholds and standard deviations as a function of SL and spatial separation in experiment 1 (n = 36)**

	Colocated		Separated	
	Low SL	High SL	Low SL	High SL
Mean thresholds	3.41	2.07	−3.62	−7.33
Standard deviation	1.33	1.51	3.04	0.353

SL, sensation level.

plots the data as a function of age, while Figure 2B, D plots the data as a function of PTA.

Thresholds were poorer in the low SL than in the high SL condition for both colocated (diamonds) and separated (circles) conditions. The mean difference between the two SL conditions was 1.34 dB in the colocated condition and 3.7 dB in the separated condition (Table 1). In the colocated condition at the high SL, shown as diamonds in Figure 2C, D, mean TMR thresholds were 0.07 dB worse than the same high SL level condition in Gallun et al. (2013). In the spatially separated condition of  $\pm 45$  degrees, shown as circles, mean TMR thresholds were 0.39 dB worse than the 2013 data. This demonstrates that the high SL condition replicated what was found in Gallun et al. (2013), and is consistent with the fact that many of the data points are the same.

The effects of age and hearing loss on thresholds for the two SL conditions are further illustrated in Table 2. In addition to reduced performance for all listeners, the relationship with hearing loss (as indicated by PTA) is also reduced at the low SL, as seen by the reduction in variance explained ( $r^2$ ) in the colocated condition from 16 to 6% and in the spatially separated condition from 35 to 16%.

A mixed-models analysis of variance (SPSS v.20) was conducted in which the within-subjects main effects of spatial separation and SL were tested in combination with the between-subjects factors of age and PTA added as covariates. The amount of variance accounted for by each of the main factors and interactions was evaluated by partial  $\eta^2$ , which expresses the amount of variance in each factor accounted for when scores are averaged across the other factors in the analysis. Note that the total amount of variance accounted for sums to 100% within a factor, but not necessarily between factors. The main effects of both spatial separation and SL were statistically significant, with spatial separation [ $F(1,33) = 181.46$ ;  $p < 0.001$ ] accounting for 85% of the variance in TMR when averaging across SL, and SL [ $F(1,33) = 21.54$ ;  $p < 0.001$ ] accounting for 40% of the variance when averaging across spatial separation. The interaction between spatial separation and SL was significant [ $F(1,33) = 5.04$ ;  $p = 0.032$ ] and accounted for 13% of the variance in TMR. In tests of between-subjects effects, PTA was significant [ $F(1,33) = 6.59$ ;  $p = 0.015$ ] and accounted for 17% of

**TABLE 2. Relationships ( $r^2$ ) between thresholds, PTA and age as a function of SL and spatial separation in experiment 1 (n = 36)**

	PTA	Colocated		Separated	
		Low SL	High SL	Low SL	High SL
Age	<b>0.203</b>	0.096	<b>0.271</b>	<b>0.115</b>	<b>0.165</b>
PTA		0.059	<b>0.156</b>	<b>0.159</b>	<b>0.353</b>

Values in bold are significant at the  $p < 0.05$  level.

PTA, pure-tone average; SL, sensation level.

the variance when averaging across SL and spatial separation, while age was not a statistically significant factor ( $p = 0.079$ ). The interaction between spatial separation and PTA was significant [ $F(1,33) = 9.43$ ;  $p = 0.004$ ] and accounted for 22% of the variance when averaging across SL. None of the other two- or three-way interactions were statistically significant ( $p$  values all greater than 0.230).

## DISCUSSION

In the colocated conditions, increasing SL resulted in slightly improved performance across ages and degrees of hearing loss. In the spatially separated condition, thresholds were considerably better for all participants when stimuli were presented at the high SL. The significant interaction found between PTA and spatial separation can be seen in Figure 2B, D, where it is clear that, regardless of SL, participants with poorer bilateral PTAs values did not achieve the amount of SRM accomplished by listeners with better PTAs. While there is a trend toward worsening performance with increasing age, the effect was not statistically significant. Gallun et al. (2013) determined that hearing loss and age are independent factors responsible for reduced SRM, but the present study only partially supports those previous findings, as age was not found to be a significant predictor of performance once PTA was taken into account, and was not shown to interact with the benefits of spatial separation.

### Experiment 2

**Examining SL and Bandwidth** • One possible explanation for the finding that PTA was a strong predictor of performance in experiment 1 is that the equating of SL across listeners did not fully restore audibility, especially for HI listeners. To continue to examine the SL effects observed in experiment 1, as well as assess the role of audibility and bandwidth on SRM, sentences were processed to (1) result in target levels of 19.5 dB SL before further processing was applied (“low SL” condition) or 39.5 dB SL (“high SL” condition), (2) equate audibility across listeners on a band-by-band basis, and (3) provide either “broadband” audibility or limited (“low pass”) audibility, similar to the audibility associated with the average sloping high-frequency hearing loss observed in our laboratory.

**Participants** • Seventy-one subjects participated (18 to 77 years; mean 43.11 years), 37 of whom had also participated in experiment 1. All participants had hearing thresholds of 70 dB HL or better at all octave frequencies between 250 and 8000 Hz, as shown in Figure 1. All had fairly symmetrical hearing at 2 kHz and below (most had differences between the ears of less than 10 dB at all frequencies, and none had differences exceeding 20 dB).

**Stimuli** • As was done in experiment 1, CRM stimuli were presented using Etymotic ER2 insert earphones and participants were asked to identify the target sentence presented at 0 degree in the presence of two maskers set at one of two spatial conditions (0 degree,  $\pm 45$  degrees). HRTFs were utilized to present the appropriate binaural cues to the listener for each spatial condition.

On each trial, before spatialization, each of the three sentences to be presented on that trial (one target and two maskers) was filtered into six component waveforms using two-octave-wide band-pass filters (first-order Butterworth using the MATLAB functions “butter” and “filtfilt”) with center frequencies of 0.5, 1, 2, 4, and 8 kHz and a low-pass filter with a cutoff frequency of 0.25 kHz (due to difficulties with unreliable outputs

from low-frequency band-pass digital filters in MATLAB). The RMS level of each of the six filtered components was then adjusted based on (1) the target level for that trial (low SL or high SL), (2) the TMR for that trial (if the components corresponded to a masker sentence) and, (3) the difference between the audiogram of the listener being tested and the audiogram of a comparison listener. For the “broadband” condition, the comparison was an “ideal” NH listener with 0 dB HL thresholds at each of the six octave frequencies. In the “low-pass” condition, however, the comparison listener had thresholds of 0 dB HL at 0.25, 0.5 and 1 kHz, 20 dB at 2 kHz, and 40 dB at 6 and 8 kHz. This listener represents a typical amount and configuration of hearing loss present in the population being tested and in the previous work from our laboratory (Gallun et al. 2013). Once the level of each component waveform had been adjusted in this manner, the six waveforms were summed together to produce the target or masker waveforms. The two-octave bandwidths of the filters ensured that all listeners would experience similar audibility, even in the case of steeply sloping audiograms. The tradeoff to this audibility is that the slope of the low-pass function was reduced relative to the target audiogram.

The goal of this band-by-band adjustment was to ensure that the audibility in each octave band was identical for all listeners, which meant sometimes attenuating the level and sometimes amplifying the level, depending on the hearing thresholds of the listener being tested and thresholds of the comparison listener

to which SL was being equated. To ensure that levels did not exceed the comfort level of the listeners, no bands were allowed to exceed RMS levels of 70 dB SPL. This resulted in reduced audibility in the high-frequency bands for 5 of the HI listeners, all of whom had PTA values greater than 28 dB HL. This only occurred at the extremes of the TMR range, however, and as such was unlikely to impact estimates of listener thresholds.

To ensure the adequacy of the processing approach, SII values were used to compare the amount of speech information available on each trial (McCreery et al. 2014). Averaged values (across all listeners) ranged from 0.026 to 0.567 and were monotonically related to TMR. Within a condition, individual listener SII values differed from mean values by an average of less than 0.001. Across conditions and TMRs, the differences from the mean SII values on average were less than 0.001 and for individual listeners were rarely greater than 0.02. The greatest deviations were for the 5 listeners with PTA values greater than 28 dB and occurred, as predicted, at the most extreme TMR values. Within the range of values associated with threshold, the values even for these listeners were within 0.02 of the mean values across listeners. These results support the use of this method for equating speech information across listeners and conditions in experiment 2.

The effects of this processing are demonstrated in the four panels of Figure 3, where the average audiogram across all listeners is plotted relative to the average right and left ear signals (after processing) for a masking sentence presented from

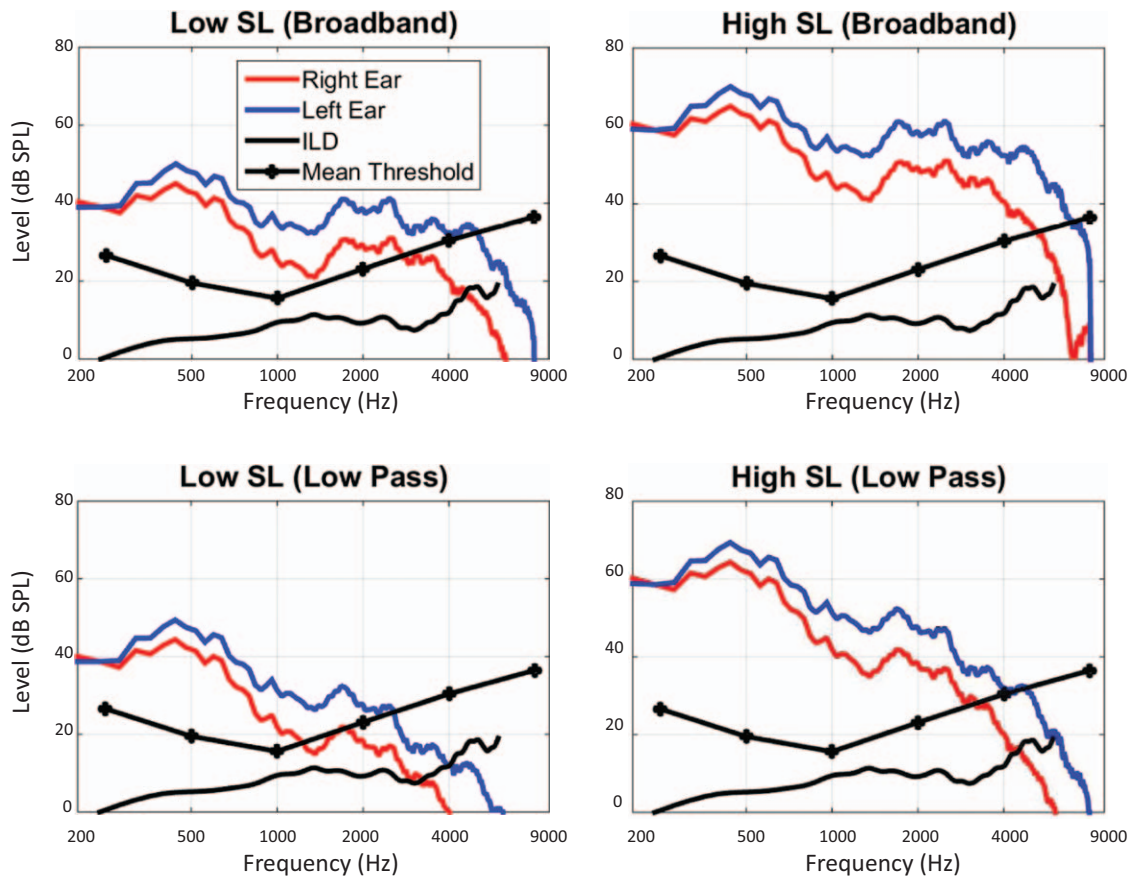


Fig. 3. Magnitude spectra (in dB SPL) of masking speech presented from the left at 45 degrees (blue lines left ear, red lines right ear); mean audiometric thresholds in dB HL (black circles); interaural level difference as a function of frequency (solid black line). Each panel represents a different combination of SL and bandwidth. The magnitude spectra represent the levels for a masker presented at a 0 dB target-to-masker ratio for a listener with the average audiogram plotted. SL indicates sensation level.

**TABLE 3.** Mean thresholds and standard deviations as a function of bandwidth (BB vs. LP), SL, and spatial separation in experiment 2 (n = 71)

	Colocated				Separated			
	Low SL		High SL		Low SL		High SL	
	LP	BB	LP	BB	LP	BB	LP	BB
Mean thresholds	3.71	2.53	2.56	2.56	-0.412	-3.44	-3.38	-4.65
Standard deviation	1.38	1.52	1.77	1.93	2.55	2.55	2.94	3.03

BB, broadband; LP, low pass; SL, sensation level.

45 degrees to the left. The average spectra at the two ears would fall between the two functions plotted for a target presented in front, and the left and right ear functions would be reversed for a masker presented from the right. The point of plotting this condition is to illustrate the ILDs, which represent the difference in dB between the two functions and are plotted at the bottom of each panel. Each panel represents a different combination of SL and bandwidth and can be used to estimate the frequency range within which audibility (and thus access to interaural cues) was maintained. Examination of Figure 3 reveals that the effective audible bandwidth, defined as the highest frequency component of the stimulus audible at both ears, is a joint function of the SL and bandwidth manipulations. Figure 3 reveals that with the manipulations employed here, increasing SL leads to a widening of the effective bandwidth, while increasing the bandwidth does not increase SL except in the region of extended bandwidth. The effective bandwidth is the lowest for the low SL condition with the low-pass processing (slightly above 1 kHz), approximately the same for the low SL broadband condition and the high SL low-pass condition (roughly 3 kHz), and highest in the high SL broadband condition (roughly 5 kHz). ILDs can be seen

to be unaffected by the signal processing, but access to the ILD cues are obviously limited by the effective bandwidth.

ITDs can change as a function of frequency, as suggested by Kuhn (1977). For the 250 Hz band, the ITD was 455  $\mu$ sec for all four combinations of bandwidth and SL, while it rose to 471  $\mu$ sec for the 500 Hz band. The ITD was 430  $\mu$ sec for the 1000 Hz band in the two broadband conditions but was 460  $\mu$ sec for the low-pass conditions. For all conditions, the ITD was 362  $\mu$ sec at 2000 Hz. For the broadband conditions, the ITD was 385  $\mu$ sec for 4000 and 8000 Hz, while for the low-pass conditions the ITD was 381  $\mu$ sec for the 4000 Hz band and 440  $\mu$ sec for the 8000 Hz band. This shows that in general the ITDs were larger in the low-frequency rather than the high-frequency region, suggesting that the low-pass conditions would still have retained important ITD cues.

It should be noted that although Figure 3 shows the levels for the average audiometric thresholds and average speech spectra, an analysis of the relative levels for each listener found that the functions were largely identical. This is consistent with the SII analysis reported above, which found that regardless of the differences in SPL across listeners, audibility was quite similar.

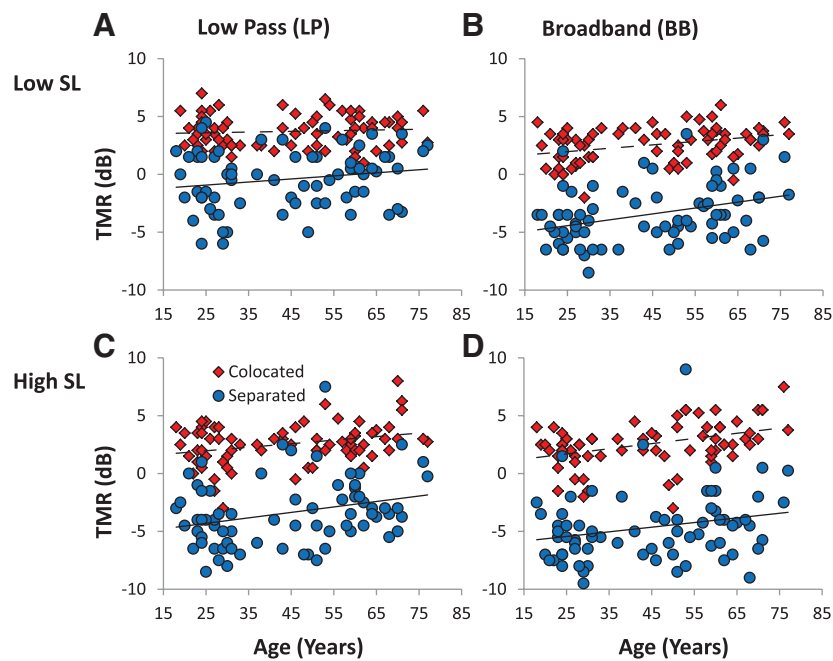


Fig. 4. Individual thresholds for the participants in experiment 2 are plotted as a function of age for the low SL condition in (A) and (B) and the high SL condition in (C) and (D). Performance with the low-pass filtered stimuli is shown in (A) and (C) and the broadband amplification is shown in (B) and (D). Dashed lines and diamonds indicate performance when target and maskers were colocated, while solid lines and circles indicate performance under conditions of spatial separation. SL indicates sensation level.

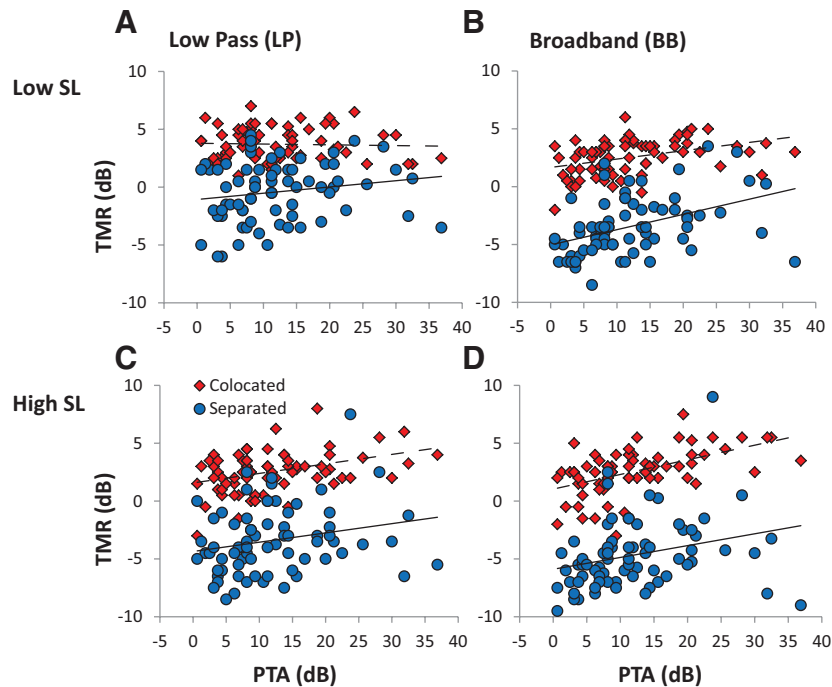


Fig. 5. Thresholds are plotted as a function of binaural PTA (A, low pass, low SL; B, broadband low SL; C, low pass, high SL; D, broadband, high SL). Dashed lines and diamonds indicate performance when target and maskers were colocated, while solid lines and circles indicate performance under conditions of spatial separation. PTA indicates pure-tone average; SL, sensation level.

**Procedures** • For experiment 2, the progressive tracking method introduced in Gallun et al. (2013) was used to determine listener thresholds for each testing condition. Target sentence levels were presented at a fixed nominal level above SRT (either the low SL or the high SL) and identical processing was applied to both targets and maskers to maintain the nominal TMRs.

The progressive tracking method consisted of presenting 20 trials, two for each of 10 TMR values starting at 10 dB TMR and decreasing in steps of 2 dB until the level of -8 dB TMR was reached. A TMR threshold estimate could be obtained by subtracting the number correct from 10 dB. This method has an accuracy of approximately 2 dB (Gallun et al. 2013) and allows for fixed TMR levels to be used as stimuli, thus permitting testing to occur via a fixed-track sound system, such as a CD player. However, for the testing purposes stated here, tracks were generated with keywords chosen randomly on each trial and presented using the computer-based processing and audio system described above.

**RESULTS**

The averaged results of experiment 2 are shown in Table 3 and the individual data are plotted with respect to age in

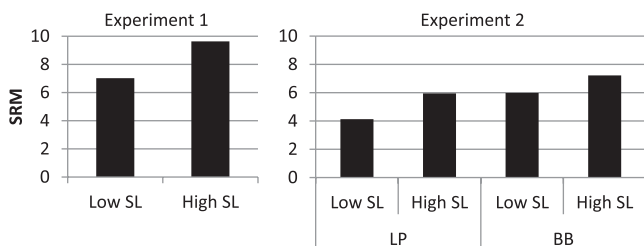


Fig. 6. Spatial release from masking plotted as a function of condition for the two experiments.

Figure 4 and with respect to PTA in Figure 5. The main effects of bandwidth, SL, and spatial separation are shown most clearly in Table 3, but are also apparent in both figures. Table 3 clearly shows how thresholds steadily improved from the colocated low SL low-pass condition on through to the spatially separated high SL broadband condition.

A mixed-models analysis of variance (SPSS v.20) was conducted in which the within-subjects main effects of spatial separation (0, ±45degrees), SL (low SL, high SL), and bandwidth (low pass, broadband) were tested in combination with the within-subjects effects of age and PTA added as covariates. Partial  $\eta^2$  was again used to examine the relative variance explained. Similar to the results of experiment 1, both the main effects of spatial separation and SL were statistically significant, with spatial separation [ $F(1,68) = 94.96; p < 0.001$ ] accounting for 58% of the variance in TMRs when averaging across SL and bandwidth, and SL [ $F(1,68) = 28.37; p < 0.001$ ] accounting for 29% of the variance when averaging across spatial separation and bandwidth. In addition, bandwidth was statistically significant [ $F(1,68) = 36.33; p < 0.001$ ] and accounted for 35% of the variance in TMRs after averaging across SL and spatial separation. The interaction between bandwidth and SL was significant [ $F(1,68) = 17.62; p < 0.001$ ] and accounted for 21% of the variance in TMRs after averaging across spatial separation.

In tests of between-subjects main effects, PTA was significant [ $F(1,68) = 4.49; p = 0.038$ ] and accounted for 6% of the variance after averaging across the other main within-subject factors. Age was not a significant predictor ( $p = 0.124$ ). The bandwidth by PTA interaction was significant [ $F(1,68) = 10.32; p < 0.002$ ] and accounted for 13% of the variance after averaging across spatial separation and SL. Only one of the possible three-way interactions was significant: between spatial separation, SL, and PTA [ $F(1,68) = 4.37; p = 0.040$ ], which accounted



**TABLE 4. Relationships ( $r^2$ ) between thresholds, PTA, and age as a function of bandwidth (BB vs. LP), sensation level, and spatial separation in experiment 2 (n = 71)**

	Colocated					Separated			
	PTA	Low SL		High SL		Low SL		High SL	
		LP	BB	LP	BB	LP	BB	LP	BB
Age	<b>0.354</b>	0.006	<b>0.114</b>	<b>0.092</b>	<b>0.165</b>	0.033	<b>0.123</b>	<b>0.08</b>	0.054
PTA		0.002	<b>0.148</b>	<b>0.143</b>	<b>0.284</b>	0.03	<b>0.177</b>	0.049	<b>0.075</b>

Values in bold are significant at the  $p < 0.05$  level.

BB, broadband; LP, low pass; PTA, pure-tone average.

for 6% of the variance after averaging across bandwidth. None of the other two-, three-, or four-way interactions were significant (all  $p$  values greater than 0.115).

This analysis shows that in the colocated conditions, only the low SL low-pass thresholds differed from the other three conditions. This suggests that once an appropriate effective audible bandwidth of roughly 3 kHz is established, increasing bandwidth either through increasing SL or amplifying the higher frequencies has little influence on colocated performance. In the spatially separated conditions, SRM was more strongly predicted by effective audible bandwidth than in the colocated conditions. Specifically, listeners performed best given stimuli with broadband filtering and a higher target SL, performed similarly with either a low SL and broadband stimulus or high SL and low-pass stimulus, and performed more poorly when the spatially separated stimuli were low-pass filtered and presented at the low SL. This same pattern is visible in Figure 6, which plots the amount of SRM for the two experiments.

The relationships between thresholds, PTA, and age shown in Table 4 are consistent with the relatively small effects found in the full statistical analysis, suggesting that when audibility is equated on an individual basis, age, and PTA have relatively small influence on performance relative to SL and spatial separation.

## GENERAL DISCUSSION

The present study found that an increase in the presented SL of target and masking stimuli improved speech recognition, at times in colocated conditions, and consistently when sound sources were spatially separated. Age did not play as large a factor in thresholds as hearing loss. Nonetheless, as was seen in experiment 1 and Gallun et al. (2013), performance generally worsened as age increased. If it is correct that the difference between the statistical results of this study and of Gallun et al. (2013) is attributable to the inclusion of participants with greater hearing loss, then it is remarkable that such small amounts of hearing loss, relative to the general population, could yield such a large change in the statistical significance of the factors of age and PTA.

Arbogast et al. (2005) tested NH and HI listeners in a spatial release task high in IM, and the amount of SRM achieved for the HI group was less than for the NH group. However, in the HI group alone, when the masker SL was raised from 32.5 to 35.1 dB, an increase in SRM was seen. This agrees with the present study's results which show that both NH listeners and HI listeners can benefit from an increase in target and masker SL.

The present study wished to examine the effects of bandwidth on SRM similar to what was done by Kidd et al. (2010) but with older and more HI listeners. Although our results showed that the general trend was for listeners with higher PTAs

to achieve poorer thresholds, there was no significant effect of age on thresholds for either of the filtered conditions.

Small increases in benefit were found to occur in the colocated low-pass conditions when the target SL was raised from the low SL to the high SL (a benefit of approximately 2dB was observed for all but the oldest listeners). This suggests that either ensuring low-frequency audibility or extending the effective bandwidth from 1 to 3 kHz, or perhaps both, can lead to improvements with even nonspatially separated speech maskers. Future study should employ different types of stimulus manipulations to more clearly examine these two effects.

With regard to the binaural cues available in the four combinations of bandwidth and SL, Figure 3 clearly shows the presence of large ILDs (5 to 10 dB) at frequencies of 1 kHz and below. It can be seen that at least nominal audibility was maintained in this frequency region for all of the stimuli. The ILD curves can also be used to estimate the better ear effect, although the use of symmetrically placed maskers essentially removes this cue (Marrone et al. 2008a). However, it is possible that the glimpsing of multiple talkers allows the better ear cue to provide brief periods of energetic or informational unmasking (Glyde et al. 2013c; Wan et al. 2014). Overall, it is difficult to determine from the data whether binaural cues or increased speech clarity improved performance.

The results of the present study have substantial implications for how amplification could improve performance in complex listening environments. The interaction between bandwidth and SL (Fig. 3) implies that increasing gain on a hearing aid can also increase access to high-frequency cues, even if the nominal bandwidth is not increased. The finding that increasing SL improved both colocated and spatially separated performance suggests that audibility enhances access to both monaural and binaural cues necessary for separating competing talkers. The small effects of age and hearing loss suggest that these benefits could be made available to a wide range of individuals who rely upon hearing aids.

## ACKNOWLEDGMENTS

We are grateful to the Veterans and non-Veterans who volunteered their time to participate in this study. We are also grateful to Meghan Stansell and Sara Sell for help with data collection and database management.

Supported by the National Institutes of Health's National Institute for Deafness and Communication Disorders (R01 DC011828) and by the VA RR&D National Center for Rehabilitative Auditory Research, which is located at the VA Portland Health Care System Medical Center. The contents of this article are the private views of the authors and should not be assumed to represent the views of the Department of Veterans Affairs or the United States Government.

The authors have no conflicts of interest to disclose.



Address for correspondence: Kasey M. Jakien, National Center for Rehabilitative Auditory Research, VA Portland Health Care System, 3710 SW US Veterans Hospital Rd., Portland, OR 97239, USA. E-mail: Kasey.Jakien@va.gov

Received October 14, 2014; accepted June 3, 2016.

## REFERENCES

- Ahlstrom, J. B., Horwitz, A. R., Dubno, J. R. (2009). Spatial benefit of bilateral hearing AIDS. *Ear Hear*, *30*, 203–218.
- Ahlstrom, J. B., Horwitz, A. R., Dubno, J. R. (2013). Spatial Separation benefit for unaided and aided listening. *Ear Hear*, *35*, 72–85.
- ANSI S3.6. (2004). *Specification for Audiometers*. New York, NY: American National Standards Institute.
- ANSI S3.5. (2007). *Methods for the Calculation of the Speech Intelligibility Index*. New York, NY: American National Standards Institute.
- Arbogast, T. L., Mason, C. R., Kidd, G. Jr. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, *117*(4 pt 1), 2169–2180.
- Best, V., Marrone, N., Mason, C. R., et al. (2012). The influence of non-spatial factors on measures of spatial release from masking. *J Acoust Soc Am*, *131*, 3103–3110.
- Best, V., Thompson, E. R., Mason, C. R., et al. (2013). An energetic limit on spatial release from masking. *J Assoc Res Otolaryngol*, *14*, 603–610.
- Bolia, R. S., Nelson, W. T., Ericson, M. A., et al. (2000). A speech corpus for multitalker communications research. *J Acoust Soc Am*, *107*, 1065–1066.
- Culling, J. F., Hawley, M. L., Litovsky, R. Y. (2004). The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. *J Acoust Soc Am*, *116*, 1057–1065.
- Dubno, J. R., Ahlstrom, J. B., Horwitz, A. R. (2002). Spectral contributions to the benefit from spatial separation of speech and noise. *J Speech Lang Hear Res*, *45*, 1297–1310.
- Gallun, F. J., Diedesch, A. C., Kampel, S. D., et al. (2013). Independent impacts of age and hearing loss on spatial release in a complex auditory environment. *Front Neurosci*, *7*, 1–11.
- Gelfand, S. A., Ross, L., Miller, S. (1988). Sentence reception in noise from one versus two sources: Effects of aging and hearing loss. *J Acoust Soc Am*, *83*, 248–256.
- Glyde, H., Buchholz, J. M., Dillon, H., et al. (2013a). The importance of interaural time differences and level differences in spatial release from masking. *J Acoust Soc Am*, *134*, EL147–EL152.
- Glyde, H., Cameron, S., Dillon, H., et al. (2013b). The effects of hearing impairment and aging on spatial processing. *Ear Hear*, *34*, 15–28.
- Glyde, H., Buchholz, J., Dillon, H., et al. (2013c). The effect of better-ear glimpsing on spatial release from masking. *J Acoust Soc Am*, *134*, 2937–2945.
- Kuhn, G. (1977). Model for the interaural time differences in the azimuthal plane. *J Acoust. Soc. Am.* *62* (1), 157–167.
- Kidd, G., Mason, C. R., Best, V., Marrone, N., (2010). Stimulus factors influencing spatial release from speech-on-speech masking. *JASA*, *128*, 1965.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J Acoust Soc Am*, *49*, Suppl 2:467+.
- Levy, S. C., Freed, D. J., Nilsson, M., et al. (2015). Extended high-frequency bandwidth improves speech reception in the presence of spatially separated masking speech. *Ear Hear*, *36*, e214–e224.
- Marrone, N., Mason, C. R., Kidd, G. Jr. (2008a). Evaluating the benefit of hearing aids in solving the cocktail party problem. *Trends Amplif*, *12*, 300–315.
- Marrone, N., Mason, C. R., Kidd, G. (2008b). Tuning in the spatial dimension: evidence from a masked speech identification task. *J Acoust Soc Am*, *124*, 1146–1158.
- Marrone, N., Mason, C. R., Kidd, G. Jr. (2008c). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *J Acoust Soc Am*, *124*, 3064–3075.
- McCreery, R. W., Brennan, M. A., Hoover, B., et al. (2014). The Influence of Audibility on Speech Recognition With Nonlinear Frequency Compression for Children and Adults With Hearing Loss. *Ear Hear*, *34*, e24–e27.
- Moore, B. C. J., Fullgrabe, C., Stone, M. A. (2010). Evaluating the benefit of hearing aids in solving the cocktail party problem. *J Acoust Soc Am*, *128*, 360–371.
- Wan, R., Durlach, N. I., Colburn, H. S. (2014). Application of a short-time version of the Equalization-Cancellation model to speech intelligibility experiments with speech maskers. *J Acoust Soc Am*, *136*, 768–776.
- Xie, B. (2013). *Head-Related Transfer Function and Virtual Auditory Display* (2nd ed.). Plantation, FL: J. Ross Publishing.