The role of interaural differences on speech intelligibility in complex multi-talker environments[a)]

Rachel L. Ellinger,[b)] Kasey M. Jakien,[c)] and Frederick J. Gallun[c),d)]
National Center for Rehabilitative Auditory Research, VA Portland Health Care System, Department of Veterans Affairs, 3710 Southwest US Veterans Hospital Road, P5-NCRAR, Portland, Oregon 97239, USA
RachelEllinger@gmail.com, Kasey.Jakien@va.gov, Frederick.Gallun@va.gov

Abstract: Interaural differences in time (ITDs) and interaural differences in level (ILDs) contribute to a listener's ability to achieve spatial release from masking (SRM), and help to improve speech intelligibility in noisy environments. In this study, the extent to which ITDs and ILDs contribute to SRM and the relationships with aging and hearing loss were examined. SRM was greatest when stimuli were presented with consistent ITD and ILD, relative to ITD or ILD alone, all of which produced greater SRM than when ITD and ILD cues were in conflict with each other. This pattern was independent of age and hearing loss.

[DDO]
Date Received: August 25, 2016 **Date Accepted:** December 18, 2016

1. Introduction

Spatial processing is very important for understanding speech in real-world environments where multiple speech signals occur simultaneously at different spatial locations. In such situations, normal hearing (NH) listeners can use spatial separation to achieve spatial release from masking (SRM), thereby improving speech intelligibility. SRM occurs, in part, because of the presence of binaural cues such as interaural differences in time (ITDs) and interaural differences in level (ILDs). While it is widely acknowledged that ITDs and ILDs contribute to SRM, the extent to which these cues are useful in isolation or in conjunction is unclear. Previous research has shown conflicting results, and much of this research has examined young normal hearing (YNH) listeners, not older listeners or individuals with hearing loss. Because hearing loss and age-related degeneration of the cochlea and central auditory system have the potential to affect the spectral and temporal resolution of the auditory system and degrade binaural cues (Glyde *et al.*, 2011), we aim to clarify the importance of ILDs and ITDs for SRM regarding individuals varying in age and hearing ability.

Glyde *et al.* (2013) found that NH listeners can achieve SRM with ILDs alone for speech-in-speech stimuli. Culling *et al.* (2004) observed that NH listeners achieve SRM with ITDs and ILDs for speech-in-speech stimuli, but that maximal release can only be achieved when both ITDs and ILDs are present. Earlier research (e.g., Hirsh, 1948; Webster, 1951; Jeffress *et al.*, 1956; Durlach, 1960, 1963; Hafter and Carrier, 1972), has accounted for much of the data on binaural release from masking for tonal signals with a range of theoretical models. All of these approaches can be characterized as based either on internal signal-processing, such as is suggested in Durlach's (1963) equalization and cancellation (EC) model, or on perceived location, such as Hafter and Carrier's (1972) lateralization model suggests.

Colburn and Durlach (1965) examined the effects of conflicting and consistent cues on the masking level difference (MLD) and found that NH listeners perform equally well with opposing ITD and ILD cues as compared to consistent cues. This was seen as support for the EC model, which proposes that ITDs and ILDs are independent and are thus unaffected by whether they are reinforcing or opposing each

[a)]The experimental work in this study was carried out as part of R.L.E.'s undergraduate thesis at Reed College.
[b)]Present address: Communication Sciences and Disorders Department, 2240 Campus Drive, Northwestern University, Evanston, IL 60208, USA.
[c)]Also at: Department of Otolaryngology/Head & Neck Surgery, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Portland, OR 97239, USA.
[d)]Author to whom correspondence should be addressed.

other. However, Hafter and Carrier (1972) found that ITD and ILD are not able to be traded completely, and reported data that were interpreted as support for a model in which detection was based on the perception of spatial separation among the auditory objects. While Domnitz and Colburn (1976) showed that both classes of models are mathematically equivalent, they also suggested that it may still be possible to distinguish between them for certain classes of stimuli.

Gallun *et al.* (2008) followed up on this possibility by exploring the role of binaural release for tone detection in multi-tone maskers or in noise with the prediction that consistent ITDs and ILDs would result in better performance for a multi-toned complex where perceived location was more likely to be the cue that listeners were using to perform the task. This hypothesis was not supported as the average results were similar for consistent and conflicting cues whether the maskers were tones or noise. However, there were large individual differences among the seven YNH listeners tested. The four listeners who performed better with the noise masker than with the multi-tone masker also performed better in the consistent condition than in the opposing condition. It was suggested that these four listeners were using perceived location to do the task, while the other three were basing their decisions on an internal signal-processing operation, such as the output of the EC model, which should provide similar results for noise and multi-tone maskers.

One of the motivations for the research presented here was to explore this question further in an experimental paradigm in which using perceived location might be more natural than in a tone-detection task. For this reason, the signals and maskers were chosen to be speech and the task was to report the words spoken rather than simply to report the presence or absence of the target. Previous work on SRM for speech has explored the use of ITDs and ILDs alone, but not whether opposing and consistent cues support the same level of performance. This work also explored the findings of Culling *et al.* (2004) and Glyde *et al.* (2013) to determine whether maximum SRM can be achieved with limited interaural cues and, because there is limited research on ITDs, ILDs, and SRM with any type of stimuli for older normal hearing (ONH), or older hearing-impaired (OHI) populations, this work aims to determine the effects of aging and hearing loss on binaural listening and SRM.

2. Methods

2.1 Listeners

Forty-six English speaking participants were recruited from the Portland metro area and the VA Portland Health Care System outpatient population via flyers, word of mouth, and from a database of past participants and were grouped based on age and hearing loss after enrollment and audiometric testing was complete. Age and audiometric information are shown by group in Table 1. Fourteen YNH listeners (age range 21–49 yr, mean age 32 yr), 17 ONH listeners (age range 51–73 yr, mean age 61.7 yr), and 15 OHI (age range 50–79 yr, mean 64.3 yr) participated. All YNH participants had bilateral speech reception thresholds (SRTs) in quiet that did not exceed 20 dB hearing level (HL), all ONH SRTs were 15 dB HL or lower, and all OHI had bilateral SRTs at or above 15 dB HL. SRT is a measure of speech reception rather than sensitivity to pure tones. As Table 1 reveals, many of the ONH listeners had substantial high-frequency loss despite all having SRTs of 15 dB HL or better. Audiometric testing with bone-conduction verified that all listeners with abnormal pure-tone thresholds at any frequency had sensorineural rather than conductive hearing loss. There were no asymmetries greater than 10 dB between right and left ear thresholds at any frequency. All participants were healthy, had no history of otological disorders, and had scores of 25 or higher on the Mini Mental State Examination (MMSE) to control for unrelated cognitive impairment (Folstein *et al.*, 1975).

2.2 Stimuli

The stimuli were sentences spoken by the first three recorded male talkers from the Coordinate Response Measure (CRM; corpus Bolia *et al.*, 2000). All CRM sentences are presented as, "Ready [CALL SIGN] go to [COLOR] [NUMBER] now." There are a total of eight possible call signs (Arrow, Baron, Charlie, Eagle, Hopper, Laker, Ringo, and Tiger), four possible colors (blue, green, red, and white), and eight possible numbers (1–8). Participants attended to the sentence containing the target call sign *Charlie* and ignored the two masking sentences containing any two of the other seven possible call signs. The target and masker CRM sentences also differed in color and number during each trial and which of the three talkers was the target varied randomly from trial to trial.

Table 1. Mean audiometric thresholds and corresponding ranges for the three listener groups.

		Left ear						Right ear					
		250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
YNH	Mean audiometric threshold (dB HL)	6.43	6.07	7.14	9.29	7.86	6.07	7.86	7.86	7.14	7.86	6.43	7.14
	Range (dB HL)	0–20	0–20	−5–20	0–25	−5–20	−5–25	5–15	5–20	0–25	−5–20	0–15	0–15
ONH	Mean audiometric threshold (dB HL)	9.12	9.41	8.82	7.94	20.88	24.71	11.18	10.00	9.12	10.00	16.18	26.18
	Range (dB HL)	0–25	0–20	0–20	−5–25	0–60	0–75	5–20	5–15	0–15	0–25	5–35	10–75
OHI	Mean audiometric threshold (dB HL)	15.33	22.33	23.67	27.33	40.00	46.67	18.00	23.33	20.33	26.00	38.00	42.00
	Range (dB HL)	5–30	10–45	10–40	5–45	10–60	20–80	10–35	10–35	10–35	10–50	10–65	10–80

Head-related-impulse-responses (HRIRs) for this experiment were from Knowles Electronics Manikin for Auditory Research (KEMAR) manikin Center for Image Processing and Integrated Computing (CIPIC) recordings acquired through the Music and Audio Research Laboratory at New York University (MARL-NYU; Andreopoulou and Roginska, 2011). HRIRs were recorded at a separation of 1 m between the microphone and the sound source. Two spatial configurations were used: co-located (all three sentences presented at 0° azimuth angle) and spatially separated (target at 0° and maskers at ±45°). For the spatially separated configuration, four conditions were compared: ITD-only, ILD-only, inconsistent, and consistent. To create the ITD-only cues, the HRIR for a single ear closest to the sound source at ±45° was copied and then shifted in the time domain by a fixed delay (Δt) of 385 μs, which is the ITD associated with a 45° separation when measured using the cross-correlation of the HRIR. This produces spectrally matched HRIRs at the two ears with only a time delay distinguishing them. It is worth noting that ITDs vary as a function of frequency (Kuhn, 1977) and, thus, the ITD used was slightly different than would occur naturally. For the ILD-only condition, both right and left ear HRIRs were used and the HRIR for the ear farthest from the sound source at ±45° was shifted in time toward the onset time of the nearer ear by 385 μs. This produced HRIRs with different spectra but nearly identical onset times (with the slight frequency variations noted above). For the inconsistent condition, changes in the time domain opposite to those applied in the ITD-only condition were applied to right and left ear HRIRs, resulting in stimuli with onset delays that were the opposite of the ILDs created by the spectral differences between the HRIRs for the two ears. For the consistent condition, the HRIRs for the right and left ear stimuli at ±45° were maintained.

2.3 Procedure

Stimuli were presented to participants via insert earphones (ER2; Etymotic Research, Elk Grove, IL) in a sound-attenuated booth at the National Center for Rehabilitative Auditory Research (NCRAR, Portland, OR). Participants selected their responses on a touch screen computer monitor. The response screen informed participants whether they had answered correctly or incorrectly following a trial, and alerted them when they had completed a set of trials. Data collection was self-paced and participants were paid for their involvement. MATLAB was used for stimulus presentation and data collection and statistical analyses were performed using SPSS Version 22 (IBM Corp, Armonk, NY). This experiment was conducted under the ethical oversight of the Reed College Institutional Review Board and the VA Portland Health Care System Institutional Review Board.

2.4 Scoring

Identification thresholds were obtained using a 1 down/1 up adaptive tracking procedure, which estimates 50% correct performance. The target level was set at 39.5 dB sensation level (SL) above the participant's SRT to ensure audibility and stayed constant throughout the experiment. Target-to-masker ratio (TMR) was varied by changing the levels of both maskers simultaneously. The starting TMR was 10 dB. Presentation levels for the maskers increased in 4 dB steps (with the TMR reducing in 4 dB steps) for the first four reversals and then changed to 2 dB steps for the remaining eight reversals. There were 12 reversals total and the last 6 reversals were averaged for the participant's thresholds. The TMR ceiling was 20 dB and the floor was −16 dB. The masker level was limited to 85 dB sound pressure level (SPL) or below in order to prevent the stimuli from being distorted or presented at a level that would be harmful to participants. This limit was imposed on the stimuli presented to 6 of the 15 OHI listeners, but all 6 were still presented with SL levels of 30 dB or above, and for 4 of the 6, the SL was 39.5. The effect of this limit was primarily to keep the maskers from exceeding a level of 85 dB, and thus the lowest TMR presented was −10 dB. The observation that none of these participants had thresholds lower than −6 dB in any condition, and four of the six had thresholds of −1 dB or above in all conditions, suggests that this limit, which ensured participant comfort, had no more than a minimal effect on audibility or estimated threshold. During the first five runs, participants went through the co-located, consistent, ITD-only, ILD-only, and inconsistent conditions in that order. Then they experienced all five conditions in a randomized order, always completing all conditions before repeating them again. Participants did this a total of 4 times, completing a total of 20 runs.

3. Results

The study was a 3×5 mixed design (group: YNH, ONH, OHI), (acoustic condition: co-located, ITD-only, ILD-only, ILD, and ITD consistent, ILD, and ITD inconsistent). The within-subjects variable was acoustic condition and the between-subjects variable was group as defined by age and hearing loss profiles. The dependent measures were TMR [shown by group and condition in Fig. 1(A)] and SRM [shown by group and condition in Fig. 1(B)]. Effect size was estimated with partial eta-squared (η^2), which varies between 0 and 1 and is approximately equivalent to the proportion of variance explained.

A mixed-model repeated measures analysis of variance (ANOVA) performed with TMR scores as the dependent variable revealed a main effect of condition [$F(4,172) = 68.552$, $p < 0.001$, $\eta^2 = 0.615$] and a main effect of group [$F(2,43) = 3.645$, $p < 0.05$, $\eta^2 = 0.145$]. There were no significant interactions between condition and group [$F(8,172) = 0.640$, not significant (ns), $\eta^2 = 0.029$]. To explore the main effect of group, pairwise comparisons were conducted among all three groups, which revealed that YNH listeners ($M_{\text{total TMR}} = -4.334$) performed significantly better than ONH ($M_{\text{total TMR}} = -1.631$) ($p < 0.05$) and OHI ($M_{\text{total TMR}} = -1.353$) listeners ($p < 0.05$), but that ONH and OHI were not significantly different ($p > 0.05$, ns). However, after Bonferroni correction for multiple comparisons, none of the group differences were statistically significant ($p > 0.05$, ns). To examine whether the lack of age and hearing loss effects were the result of using a group analysis, the ANOVA was repeated with age and hearing loss entered as covariates. The patterns of results were largely similar,

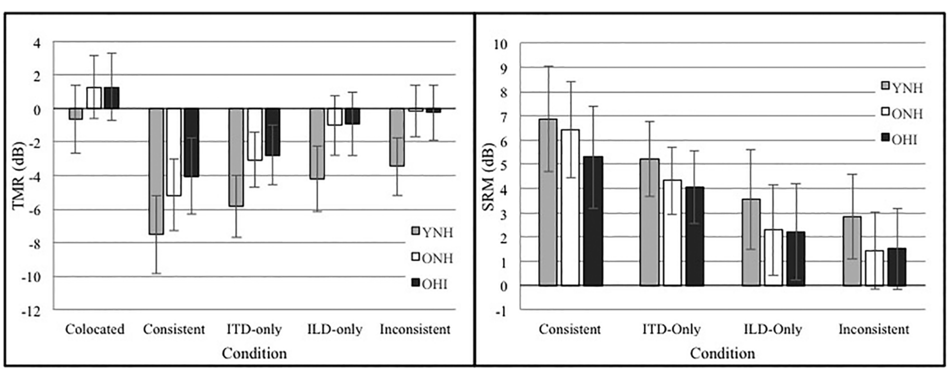

Fig. 1. (A) TMRs (dB) plotted for each condition for the three listener groups. (B) SRM (dB) plotted for each condition for the three listener groups. In both (A) and (B), error bars are ±1 standard error of the mean.

even when various measures of pure-tone sensitivity (both for high and low frequencies) were used as the measure of hearing loss.

To explore the main effect of condition, pairwise comparisons were conducted that revealed that all five conditions were significantly different from each other ($p < 0.005$). After Bonferroni correction for multiple comparisons, however, the difference between the ILD-only condition and the inconsistent condition was no longer significant ($p > 0.05$, ns). Significant differences remained for all other comparisons. Figure 1(A) shows that the observed order of average TMR thresholds from highest to lowest was: co-located ($M_{\text{total TMR}} = 0.628$) > inconsistent ($M_{\text{total TMR}} = -1.299$) > ILD-only ($M_{\text{total TMR}} = -2.053$) > ITD-only ($M_{\text{total TMR}} = -3.903$) > consistent ($M_{\text{total TMR}} = -5.570$).

SRM was calculated by measuring the difference in TMR between each separated condition and the co-located condition to reduce the variance across listeners unrelated to the use of binaural cues. A mixed-model repeated measures ANOVA performed with SRM as the dependent measure revealed a main effect of condition [$F(3,129) = 79.914$, $p < 0.005$, $\eta^2 = 0.650$]. There was no significant main effect of group [$F(2,43) = 0.649$, ns, $\eta^2 = 0.029$] and no significant interaction between condition and group [$F(6,129) = 0.626$, ns, $\eta^2 = 0.028$]. A set of pairwise comparisons revealed that all conditions were significantly different from each other even after Bonferroni correction for multiple comparisons ($p < 0.05$). Overall, the greatest amount of SRM was achieved by all three groups in the consistent condition ($M_{\text{total SRM}} = 6.193$), with only slightly less SRM obtained in the ITD-only condition ($M_{\text{total SRM}} = 4.507$). Less SRM was obtained in the ILD-only condition ($M_{\text{total SRM}} = 2.645$), and the least amount of SRM was achieved in the inconsistent condition ($M_{\text{total SRM}} = 1.886$).

Pearson correlations were run to assess the degree to which age and hearing loss (as represented by bilateral SRTs) were associated with TMR and SRM across conditions. To correct for multiple comparisons, significance was evaluated at a *p*-value of 0.005, based on the Bonferroni inequality. Significant positive correlations were observed between TMR and age in all conditions (R^2 values between 0.10 and 0.23; $p < 0.005$) except for the co-located condition ($R^2 = 0.10$; $0.05 > p > 0.005$, ns). TMR and hearing loss were significantly positively correlated for consistent ($R^2 = 0.21$; $p = 0.001$), but were not significantly correlated for any of the spatially separated conditions after correcting for multiple comparisons (R^2 values between 0.10 and 0.21; $0.05 > p > 0.005$, ns). In the co-located condition the correlation with hearing loss was not significant even before correction ($R^2 < 0.01$; $p > 0.05$, ns). This suggests that age was a better predictor of TMR across conditions than was hearing loss, which is consistent with the finding that thresholds for the ONH and OHI groups were similar to each other and different from the YNH thresholds. One potential reason for this may be the correlation of age and high-frequency hearing loss observable in Table 1. Repeating the correlational analyses with high-frequency thresholds (average of 2, 4, and 8 kHz) rather than SRT as the measure of hearing loss had no substantial impact on the patterns of performance.

Significant negative correlation was observed between SRM and hearing loss for the consistent condition ($R^2 = 0.19$; $p < 0.005$), but not for ITD-only or inconsistent (R^2 values between 0.06 and 0.08; $p > 0.05$, ns). Correlations with SRM and hearing loss were not significant in the ILD-only condition after correction ($R^2 = 0.09$; $0.05 > p > 0.005$, ns). No significant correlations were observed between age and SRM ($R^2 < 0.09$; $p > 0.05$, ns). These results suggest that SRM is less affected by age and hearing loss than is TMR. Repeating these analyses with pure-tone sensitivity (to both high and low frequencies) as the measure of hearing loss had no significant effect on the strength of the relationships.

4. Discussion and conclusion

The present study investigated the effects of hearing loss, aging, and binaural differences on TMR and SRM for speech stimuli with a multi-talker speech masker. The significant effect of group and non-significant difference in TMR between ONH and OHI listeners indicates that aging (and/or high-frequency hearing loss) affects speech intelligibility when binaural cues are limited. This is also consistent with the significant correlations between age and TMR. The non-significant interaction between group and condition in the TMR ANOVA shows that the pattern of performance for the three groups was not statistically different, suggesting similar effects of binaural differences despite the reduced overall performance for the older listeners. The SRM ANOVA did not show a group main effect or interaction of group and condition, however, suggesting that while TMR increased (indicating worse performance), SRM was largely unaffected. The only hint of an effect on SRM was the significant negative correlation of SRT and SRM in the consistent condition.

These results are a surprising contrast with the results of Glyde *et al.* (2013), who found that maximum SRM can be achieved with ILDs alone. The contrast with results of Glyde *et al.* (2013) may reflect different speech materials and/or different stimulus processing, as well as the use of a 90° spatial separation, which produces much greater differences in ILD than ITD as compared with the 45° difference used in this study. While it is also the case that this experiment examined the effects of aging and hearing loss as well, the main findings were observable in the YNH group, suggesting that the range of ages and hearing losses was not responsible for the contrasting findings.

The results of this experiment showed that significant SRM can be achieved with ITDs or ILDs alone, even for those with substantially reduced performance overall, and that maximum SRM is achieved when both cues are preserved, building on the findings of Gallun *et al.* (2008) and Culling *et al.* (2004). Gallun *et al.* (2008) and Colburn and Durlach (1965) both found similar average binaural masking level differences (BMLDs) for consistent and inconsistent ITD and ILD conditions when the target was a tone in a noise masker. In this study, significant differences between reinforcing and conflicting conditions were found. We therefore conclude that these results support the hypothesis that both older and younger listeners rely both on internal signal-processing operations that make target speech more detectable and on perceived differences in spatial location to focus attention on a speech target in the presence of similar speech maskers.

The lack of age and hearing loss effects on SRM are somewhat surprising given the previous literature, but are probably largely related to the large within-group variability [see the error bars in Fig. 1(B)]. One potential explanation is that there is a factor (or set of factors) even more important than age and mild to moderate hearing loss, and that the strength of this factor varies across all groups. Factors such as working memory, cognitive load, cognitive ability, temporal processing ability, and attention may play important roles either as independent factors or through interacting with more basic auditory sensitivity. Consequently, age and audiometry alone may provide too gross of a measure to accurately assess a listener's ability to use binaural cues in complex listening conditions. In conclusion, this study found that listener performance and SRM are affected by which binaural cues are present, with the greatest speech intelligibility occurring when listeners are presented with consistent cues, followed by ITD-only cues, ILD-only cues, inconsistent cues, and co-located cues. These results support the hypothesis that listeners trying to understand speech in multi-talker environments make use of differences in perceived locations to focus on the target talker, rather than using internal signal-processing to cancel the maskers in a manner that treats ILD and ITD as independent factors. As listeners age (and as hearing loss increases), the ability to understand a target sentence amidst masking speech becomes poorer, as evidenced by reduced TMR across all conditions. Remarkably, however, for the older listeners tested here, the pattern of performance and amount of SRM achieved remained relatively unchanged.

Acknowledgments

Thank you to all of the participants who volunteered their time to be involved in this experiment. Thank you to Sean Kampel, Meghan Stansell, and Nirmal Srinivasan for help with data collection, Kathryn Oleson for help with data analysis, and Enriqueta Canseco-Gonzalez for advisement while working on this experiment and the thesis document. This work was supported by National Institutes of Health/National Institute on Deafness and Other Communication Disorders (NIH/NIDCD; R01 DC011828), Department of Veterans Affairs Rehabilitation Research & Development Service (VA RR&D), and Reed College Department of Psychology. 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.

References and links

Andreopoulou, A., and Roginska, A. (**2011**). "Towards the creation of a standardized HRTF repository," in *131st Audio Engineering Society Conference*, New York, NY, pp. 1–6.

Bolia, R. S., Nelson, W. T., and Ericson, M. A. (**2000**). "A speech corpus for multitalker communications research," J. Acoust. Soc. Am. **107**(2), 1065–1066.

Colburn, H. S., and Durlach, N. I. (**1965**). "Time-intensity relations in binaural unmasking," J. Acoust. Soc. Am. **38**, 93–103.

Culling, J. F., Hawley, M. L., and 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**(2), 1057–1065.

Domnitz, R. H., and Colburn, H. S. (**1976**). "Analysis of binaural detection models for dependence on interaural target parameters," J. Acoust. Soc. Am. **59**(3), 598–601.

Durlach, N. I. (**1960**). "Note on the equalization and cancellation theory of binaural masking level differences," J. Acoust. Soc. Am. **32**(8), 1075–1076.

Durlach, N. I. (**1963**). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. **35**(8), 1206–1218.

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (**1975**). "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician," J. Psych. Res. **12**(3), 189–198.

Gallun, F. J., Durlach, N. I., Colburn, S., Shinn-Cunningham, B. G., Best, V., Mason, C. R., and Kidd, G. (**2008**). "The extent to which a position-based explanation accounts for binaural release from informational masking," J. Acoust. Soc. Am. **124**(1), 439–449.

Glyde, H., Buchholz, J. M., Dillon, H., Cameron, S., and Hickson, L. (**2013**). "The importance of interaural time differences and level differences in spatial release from masking," J. Acoust. Soc. Am. **134**(2), EL147–EL152.

Glyde, H., Hickson, L., Cameron, S., and Dillon, H. (**2011**). "Problems hearing in noise in older adults: A review of spatial processing disorder," Trends Amplif. **15**(3), 116–126.

Hafter, E. R., and Carrier, S. C. (**1972**). "Binaural interaction in low-frequency stimuli: The inability to trade time and intensity completely," J. Acoust. Soc. Am. **51**(6), 1852–1862.

Hirsh, I. J. (**1948**). "The influence of interaural phase on interaural summation and inhibition," J. Acoust. Soc. Am. **20**, 536–544.

Jeffress, L. A., Blodgett, H. C., Sandel, T. T., and Wood, C. L. I. (**1956**). "Masking of tonal signals," J. Acoust. Soc. Am. **28**(3), 416–426.

Kuhn, G. F. (**1977**). "Model for the interaural time differences in the azimuthal plane," J. Acoust. Soc. Am. **62**(1), 157–167.

Webster, F. A. (**1951**). "The influence of interaural phase on masked thresholds I. The role of interaural time-deviation," J. Acoust. Soc. Am. **23**, 452–462.