Research Article

An Energetic Limit on Spatial Release from Masking

VIRGINIA BEST,^{1,2} ERIC R. THOMPSON,^{1,3} CHRISTINE R. MASON,¹ AND GERALD KIDD JR¹

¹*Department of Speech, Language and Hearing Sciences and Hearing Research Center, Boston University, Boston, MA 02215, USA*

²*National Acoustic Laboratories, 16 University Ave, Macquarie University, Sydney, NSW 2019, Australia*

³*Ball Aerospace & Technologies Corporation, Fairborn, OH 45324, USA*

Received: 13 December 2012; Accepted: 17 April 2013; Online publication: 7 May 2013

ABSTRACT

This study tested the hypothesis that energetic masking limits the benefits obtained from spatial separation in multiple-talker listening situations, particularly for listeners with sensorineural hearing loss. A speech target was presented simultaneously with two or four speech maskers. The target was always presented diotically, and the maskers were either presented diotically or dichotically. In dichotic configurations, the maskers were symmetrically placed by introducing interaural time differences (ITDs) or infinitely large interaural level differences (ILDs; monaural presentation). Target-to-masker ratios for 50 % correct were estimated. Thresholds in all separated conditions were poorer in listeners with hearing loss than listeners with normal hearing. Moreover, for a given listener, thresholds were similar for conditions with the same number of talkers per ear (e.g., ILD with four talkers equivalent to ITD with two talkers) and hence the same energetic masking. The results are consistent with the idea that increased energetic masking, rather than a specific spatial deficit, may limit performance for hearing-impaired listeners in spatialized speech mixtures.

Keywords: hearing loss, spatial unmasking, binaural

INTRODUCTION

At least two distinct kinds of auditory masking can degrade speech intelligibility in situations containing competing talkers. ‘Energetic masking (EM)’ occurs when the competing talkers overlap acoustically with the target and render it inaudible. ‘Informational masking (IM)’ describes interference occurring more centrally, and is often attributed to confusion between audible target and masker elements (Kidd et al. 2008). When competing talkers are presented from the same location, the dominant kind of interference appears to be IM (Brungart et al. 2006, 2009). However, in two recent studies by Best and colleagues (Best et al. 2012, 2013) it was observed that for spatially separated talkers, thresholds for a given listener do not vary much when the IM content of the masker is varied (i.e., when a speech masker is time-reversed or replaced with noise). This suggests that IM may be largely eliminated under these conditions, leaving primarily EM (see also Freyman et al. 2007).

In numerous studies, listeners with sensorineural hearing impairment (HI) have been shown to receive less benefit from the spatial separation of competing talkers than listeners with normal hearing (NH). For example, when following a frontal target talker in the presence of two interfering talkers, NH listeners demonstrate a large spatial release from masking (up to 20 dB) when the interferers are separated symmetrically to either side, but this value is substantially reduced by hearing loss (Marrone et al. 2008; Best et al. 2012; Spencer et al. 2012; Glyde et al. 2013). For the symmetric talker paradigm, it has been noted several times that spatial benefits are reduced as a result of higher thresholds in the *separated* configuration; *co-located* thresholds tend to be similar across

Portions of this work were presented at the Association for Research in Otolaryngology MidWinter Meeting, San Diego, February 2012.

Correspondence to: Virginia Best · National Acoustic Laboratories · 16 University Ave, Macquarie University, Sydney, NSW 2019, Australia. Virginia.Best@nal.gov.au

listener groups. This apparent deficit in spatial conditions has motivated the idea that spatial processing is at the heart of the problem. However, another interpretation is that hearing loss is detrimental in conditions dominated by EM, and not those dominated by IM. According to the above description, this would appear as a deficit in separated configurations that is not revealed in co-located configurations.

This interpretation is consistent with several experiments that have isolated IM from EM and concluded that NH and HI listeners are equally susceptible to IM but more susceptible to EM (e.g., Helfer and Freyman 2008; Agus et al. 2009). Note that increased susceptibility to EM is equivalent to increased susceptibility to noise, which is associated with hearing loss for a variety of reasons including inaudibility of high-frequency speech components, reduced sensation level, and poor frequency selectivity (Moore 2007; Dillon 2012).

We propose that speech intelligibility in multitalker mixtures is limited by EM, which is determined both by the acoustic characteristics of the stimulus (i.e., how much masker energy is present) and by characteristics of the listener (i.e., how susceptible he/she is to noise). This limit ultimately determines how much release from IM is possible when competing speech signals are spatially separated, as it effectively determines the headroom for improvement.

The current experiment tested two specific hypotheses. Our first hypothesis was that the amount of EM in multitalker mixtures imposes a limit on speech intelligibility. We varied the EM systematically by varying the number of masker talkers. Apart from the extra energy added, increasing the number of talkers tends to ‘fill in’ the spectrotemporal gaps in the stimulus, leaving fewer opportunities to obtain clean glimpses of the target. In the limit, of course, adding independent talkers together results in speech-like noise, which is an optimal energetic masker (Brungart 2001; Brungart et al. 2001). To obtain a more direct measure of the amount of EM in each configuration, a noise condition was included, in which the speech maskers were replaced with spectrotemporally matched noises. In addition to varying the number of maskers, both interaural time and interaural level differences (ITDs and ILDs) were used separately to spatialize them. This gave us another level of control over the amount of EM, since separating the maskers via ITD leaves masker energy in both ears, whereas separating them via ILDs removes masker energy from one ear. We measured speech reception thresholds and predicted that each listener’s performance would be systematically related to EM *at the ear* in the different spatialized conditions, regardless of the masker type, number, or arrangement.

Our second hypothesis was that for a given stimulus, the effective EM is higher in HI listeners than in NH listeners, leading to a more restrictive performance limit. Thus, we predicted that HI listeners would perform more poorly in all spatialized conditions, as a result of an individual EM limit related to their audiometric profile. Importantly, if a monaural EM-based limit could account for performance across both ITD and ILD conditions it would provide evidence against the existence of a specific binaural deficit causing poor performance on this task in HI listeners.

METHODS

Participants

Participants were 11 NH listeners (18–33 years; mean 23 years) and 11 HI listeners with bilateral symmetric sensorineural hearing loss (18–43 years; mean 25 years). Audiograms for the HI group are shown in Figure 1. Listeners were paid for their participation and gave informed consent. All procedures were approved by the Boston University Institutional Review Board.

Stimuli and procedures

Digital stimuli were generated on a PC using MATLAB software, D/A converted and attenuated using Tucker-Davis Technologies hardware (System II), and presented over Sennheiser headphones (HD 280 Pro). Listeners were seated in a sound-treated booth fitted with a monitor and mouse, and indicated their responses by clicking with the mouse on a graphical user interface.

Target stimuli were taken from the coordinate response measure (CRM) corpus (Bolia et al. 2000). Sentences in this corpus have the form “Ready < callsign > go to < color > < number > now,” with

FIG. 1. Audiograms for the individual HI listeners (averaged across left and right ears, gray lines). Black symbols and error bars show the mean and standard deviation across the group.

eight/four/eight alternatives for the callsign/color/number keywords respectively (e.g., “Ready Charlie go to blue six now”). Only the four male talkers in the corpus were used, and the target was always given the callsign “Baron.”

Both speech and noise maskers were used. Speech maskers consisted of two or four talkers (different from the target talker) speaking CRM sentences. The keywords spoken by the masker talkers were different from the target and, where possible, different from each other (as there are only four colors in the corpus, two of the maskers in the four-talker masker always uttered the same color). Noise maskers were generated by first selecting two or four speech maskers as just described, then extracting the broadband envelope from each talker using a Hilbert transform, low-passing it at 50 Hz, and using it to modulate a noise with a spectrum matched to the long-term average spectrum of the male voices in the CRM corpus. Note that even though the different sentences in the CRM corpus have the same overall structure, the different talkers have different speaking rates and rhythms, which means that the envelopes are not tightly correlated across sentences.

Individualized linear gain (according to the NAL-RP prescription; Byrne et al. 1990) was applied for HI listeners before presentation over headphones. Each masker talker was fixed at 60 dB SPL (pre-gain), and the target level was varied adaptively using a one-up one-down staircase procedure to track the target-to-masker ratio (TMR) giving 50 % correct performance. No feedback was provided during testing.

Experimental conditions

Figure 2 shows a schematic of the different experimental conditions. The target was always presented diotically, such that it was perceived in the center of

the head. In the diotic conditions, all maskers were also presented diotically (DIO2 and DIO4), which is analogous to the typical co-located configuration of many free-field studies. In the dichotic conditions, ILDs or ITDs were used to move the perceived location of the masker talkers to either side of the head (akin to a symmetrically separated configuration in the free field). In the ILD conditions (ILD2 and ILD4), each masker was presented only to one ear (i.e., with an infinite ILD), and the maskers were split such that each ear received one of the two (or two of the four) maskers (as per Pollack and Pickett 1958). In the ITD conditions (ITD2 and ITD4), each masker was lateralized to one side using a simple time delay of 0.6 ms. Again, the maskers were split such that half of the maskers were perceived on each side of the head, but in this case, energy from all maskers was present in each ear. The key feature of this design is that there are two conditions with the same EM by our definition (ILD4 and ITD2, each having two masker talkers per ear), despite those two conditions having a different number of talkers in total and different spatial cues imposed.

The 12 masked conditions (speech/noise, two/four maskers, diotic/ITD/ILD) were tested in a random order within an experimental session. Each session began with a quiet (target-only) condition, giving 13 adaptive tracks in total, and took around 45 min to complete. Four sessions were completed by each participant over several visits.

RESULTS

Quiet performance

Quiet speech reception thresholds were considerably lower in the NH group (mean 19 dB SPL) than in the HI group (mean 42 dB SPL). Moreover, quiet

FIG. 2. Schematic of the different experimental conditions, depicting the target (T) and maskers (M1-M4). In ILD conditions, maskers are presented to one ear only (i.e., infinite ILDs). In ITD conditions, maskers are presented to both ears and underlining indicates which ear leads in time. The arrow indicates the two conditions that have the same number of maskers in each ear.

thresholds in the HI group were correlated with audiometric thresholds (average of 0.5, 1, and 2 kHz; $r=0.77$, $P=0.006$). This was expected given that NAL-RP is not designed to restore full audibility, but it does mean that target and masker stimuli were presented at a lower overall sensation level to HI listeners.

Masked performance

Figure 3 shows threshold TMRs for speech and noise maskers (top and bottom panels) under each of the experimental conditions. Mean thresholds for NH and HI groups are shown in black and white, respectively. Figure 4 shows the amount of spatial release from masking for each condition and each group, obtained by simply subtracting dichotic from diotic thresholds.

For speech maskers (Fig. 3 top panel), diotic thresholds were high (positive TMRs) and almost identical across groups, consistent with this condition being dominated by IM. Thresholds were lower in the dichotic conditions, suggesting a release from IM, but did not get as low in the HI group as in the NH group. A mixed ANOVA revealed significant effects of condition ($F_{5,100}=146.6$, $P<0.001$), group ($F_{1,20}=18.2$,

FIG. 3. Mean target-to-masker ratios at threshold (expressed per masker) for NH and HI listeners (filled and open symbols, respectively). Error bars show across-subject standard deviations. Top panel: speech maskers, bottom panel: noise maskers.

FIG. 4. Mean spatial benefits for NH and HI listeners (filled and open symbols, respectively). Error bars show across-subject standard deviations. Top panel: speech maskers, bottom panel: noise maskers.

$P<0.001$), and a significant interaction ($F_{5,100}=19.9$, $P<0.001$). Planned comparisons (t tests, $P<0.05$) confirmed that means were different between groups for the four dichotic conditions but not the two diotic conditions.

For noise maskers (Fig. 3 bottom panel), diotic thresholds were lower than for speech maskers, consistent with there being little or no IM. Moreover, the HI group showed poorer thresholds on average for both diotic and dichotic conditions. A mixed ANOVA revealed significant effects of condition ($F_{5,100}=348.6$, $P<0.001$), group ($F_{1,20}=28.5$, $P<0.001$), and a significant interaction ($F_{5,100}=22.8$, $P<0.001$). Planned comparisons (t tests, $P<0.05$) confirmed that means were different between groups for all six conditions.

A qualitative examination of the mean dichotic thresholds in Figure 3 for both speech and noise maskers suggests that performance was systematically related to EM. EM was determined by the number of masker talkers (or noises) at the ear, which could be either one (ILD2), two (ILD4 and ITD2), or four (ITD4). The ranking of thresholds across the four conditions was related directly to this number; thresholds were essentially identical within each group for the two conditions with the same number of masker

talkers in each ear (ILD4 and ITD2), were lower in the condition with fewer talkers in each ear (ILD2), and were higher in the condition with more talkers in each ear (ITD4).

Spatial benefit

Spatial benefit (the difference between diotic and dichotic thresholds) was calculated for all conditions and is shown in Figure 4. For speech maskers, spatial benefits were larger in the NH group than in the HI group for all conditions (Fig. 4 top panel; difference between groups ranges from 3.5 to 9.6 dB). A mixed ANOVA revealed significant effects of condition ($F_{3,60}=88.5$, $P<0.001$), group ($F_{1,20}=22.9$, $P<0.001$), and a significant interaction ($F_{3,60}=13.4$, $P<0.001$). Planned comparisons (t tests, $P<0.05$) confirmed that means were different between groups for all four conditions.

For noise maskers, spatial benefits were smaller on average than for speech maskers, and the difference between groups was reduced (Fig. 4 bottom panel; difference between groups ranges from 0.8 to 3.8 dB). A mixed ANOVA revealed significant effects of condition ($F_{3,60}=140.5$, $P<0.001$), group ($F_{1,20}=23.4$, $P<0.001$), and a significant interaction ($F_{3,60}=12.3$, $P<0.001$). Planned comparisons (t tests, $P<0.05$) confirmed that means were different between groups for all conditions but ITD4.

SNR analysis

To examine the relationship between performance and EM more quantitatively, we expressed the separated thresholds for each listener with reference to the overall masker energy at the ear rather than the energy of each individual masker at the “source” (i.e., converted from TMR to SNR). This adjustment accounted for the drop in SNR caused by additional talkers at the ear. As a first approximation, we simply calculated the drop in the long-term broadband SNR using the root-mean-square of the signals presented to the ear under the different dichotic conditions. Hence, the thresholds for ILD2 were unchanged (as there was one masker talker in each ear), thresholds for ILD4/ITD2 were reduced by 3 dB, and thresholds for ITD4 were reduced by 6 dB.

Figure 5 shows these SNRs at threshold for each listener (gray lines) as well as the mean across listeners (black lines and symbols). If our hypothesis is true that each listener has an individual EM limit and if broadband SNR at the ear is sufficient to capture EM, then thresholds should be consistent across spatial configurations when plotted in this way (i.e., the lines joining the four conditions should be horizontal). For noise maskers (Fig. 5 bottom panels), the SNR functions were generally quite flat, especially for the

HI group. A mixed ANOVA conducted on the SNR values for noise maskers confirmed a significant interaction between spatial condition and group ($F_{3,60}=27.3$, $P<0.001$). Individual ANOVAs for each group showed that the effect of spatial condition was significant for the NH group ($F_{3,30}=32.3$, $P<0.001$) but not the HI group ($F_{3,30}=2.6$, $P=0.07$). For speech maskers (Fig. 5 upper panels), many HI listeners again showed relatively flat functions, but the majority of NH listeners did not. A mixed ANOVA conducted on the SNR values for speech maskers confirmed a significant interaction between spatial condition and group ($F_{3,60}=27.3$, $P<0.001$). Individual ANOVAs for each group showed that the effect of spatial condition was significant for the NH group ($F_{3,30}=25.2$, $P<0.001$), and the small effect of spatial condition for the HI group also reached significance (upper right panel, $F_{3,30}=3.5$, $P=0.03$).

For the NH group, SNRs at threshold for both noise and speech maskers clearly deviated from the “fixed SNR” prediction. Instead they showed an increasing trend overall, being lowest for the ILD2 condition and highest for the ITD4 condition (but equivalent for ILD4 and ITD2, as expected). The deviations were especially large for some listeners in the speech masker condition. Interestingly, the two NH listeners with the poorest thresholds in this condition show relatively flat functions that are comparable to the HI functions. The inability of SNR to capture the NH results may be a result of the simplistic calculation of SNR used here, which does not take into account the opportunity for “glimpsing” in the different speech mixtures. For example, in the ILD2 condition, where there is only one masker in each ear, there are many more opportunities for catching clean glimpses of the target than in the spectrotemporally denser ILD4/ITD2 conditions. If NH listeners are better able to make use of these brief moments where the target is unmasked than HI listeners (e.g., Festen and Plomp 1990; George et al. 2006; Hall et al. 2012) then a broadband, long-term metric would be expected to fall short. Indeed Brungart et al. (2009) show that NH listeners experience more EM when the number of maskers is increased even if the overall SNR is held constant.

For HI listeners, we predicted that individual EM limits would be related to individual audiometric profiles. To examine this, an estimate of the “SNR limit” for each listener was calculated by averaging his/her transformed thresholds for the four dichotic conditions (i.e., averaging across the gray lines in the right panels of Fig. 5). Figure 6 shows the estimated SNR limit plotted against pure tone average (the average of left and right ear thresholds at 0.5, 1, and 2 kHz) for each HI listener. The correlation between

FIG. 5. Individual dichotic thresholds expressed in terms of the broadband signal-to-noise ratio at the ear (gray lines). Black symbols and error bars show the mean and standard deviation across the group. Top row: speech maskers, bottom row: noise maskers. Left column: NH listeners, right column: HI listeners.

the two measures was highly significant for both speech maskers (circles; $r=0.81$, $P=0.002$) and noise stimuli (squares; $r=0.79$, $P=0.004$). While a strong correspondence was found between the slopes of the least squares fits for speech (solid lines; slope=0.22 dB/dB) and noise (dashed lines; slope=0.21 dB/dB), SNR limits were consistently higher for speech

maskers (by approximately 4 dB). This offset may reflect residual IM with speech maskers, which simply increases the SNR needed to perform at a given level.

DISCUSSION

We examined spatial release from masking in various multitalker mixtures. While all listeners demonstrated a release from masking under dichotic conditions, the benefit was reduced in HI listeners. These general findings are consistent with previous studies employing similar materials and configurations (Marrone et al. 2008; Best et al. 2012; Spencer et al. 2012; Glyde et al. 2013).

Our first hypothesis was that the amount of EM in multitalker mixtures imposes a limit on speech intelligibility. We predicted that performance in the different spatialized conditions would be systematically related to EM, defined simply by the number of masker talkers at the listener's ear. This was found to be the case, with the strongest demonstration coming from the fact that thresholds were identical for the two conditions with the same number of masker talkers in each ear (ILD4 and ITD2), despite those two conditions having a different number of talkers in total, and different spatial cues imposed. A similar pattern of results was also observed when the maskers were noise stimuli, which further supports an explanation based on EM. On the whole, these findings

FIG. 6. Estimated SNR limits for each HI listener as a function of pure-tone average for speech maskers (circles) and noise maskers (squares). Solid and dashed lines show least-squares fits to the speech and noise data, respectively.

strongly suggest that *monaural* factors limited performance in the spatialized mixtures used in this experiment.

The striking equivalence of the thresholds for ILD4 and ITD2 is somewhat surprising if one considers the well-known ability of the auditory system to gain a speech intelligibility advantage from discrepancies in interaural phase between targets and maskers (the binaural intelligibility level difference; e.g., Licklider 1948; Levitt and Rabiner 1967). Because these binaural mechanisms are most efficient for interaurally correlated maskers (even in the case of multiple maskers, e.g., Culling et al. 2004), one might predict better thresholds in the ITD2 condition (which contains correlated maskers) compared to the ILD4 condition (in which the maskers are uncorrelated). It may be that ILD4 condition, due to the infinite ILDs, provides more opportunities to listen in the gaps at different times in each ear, offsetting any BMLD-related advantage afforded by the ITD2 condition.

Our second hypothesis was that listeners with sensorineural hearing loss would exhibit more EM than NH listeners, and that this would in turn reduce their spatial release from masking. Indeed, we found that dichotic thresholds for HI listeners were poorer across all conditions, regardless of the masker number or spatial cue imposed. Moreover, their thresholds were well-explained in terms of a fixed “SNR limit” that was related to the severity of their hearing loss. This provides a plausible explanation for the reduced benefit of spatial separation in this population, based on monaural factors alone, without the need to consider specific deficits in the processing of ITD or ILD. This interpretation could explain why attempts to relate reduced spatial benefits in the HI population to specific deficits in binaural or localization abilities have been largely unsuccessful (e.g., Noble et al. 1997; Spencer et al. 2012), while spatial benefits are strongly correlated to *monaural* measures such as audiometric thresholds (e.g., Marrone et al. 2008; Glyde et al. 2013).

Further investigations will be required to assess whether or not EM limits generalize to other situations. For example, the spatial differences imposed were rather extreme, and it is possible that EM limits would not be reached if smaller separations (i.e., smaller ITDs/ILDs) had been used. It is also possible that our young healthy listeners represent ‘ideal’ cases who reach their EM limit only because they have no other performance-limiting deficits; this might not be the case for older listeners, or listeners with central spatial processing deficits (e.g., Dubno et al. 2008; Glyde et al. 2011).

CONCLUSIONS

The results are consistent with a performance limit in multitalker mixtures that is related to energetic masking.

This limit may at least in part explain the reduced benefit from cues such as spatial separation often observed in listeners with sensorineural hearing loss.

ACKNOWLEDGMENTS

This work was supported by grants from NIH/NIDCD and AFOSR.

REFERENCES

- AGUS TR, AKEROYD MA, GATEHOUSE S, WARDEN D (2009) Informational masking in young and elderly listeners for speech masked by simultaneous speech and noise. *J Acoust Soc Am* 126:1926–1940
- BEST V, THOMPSON ER, MASON CR, KIDD G JR. (2013) Spatial release from masking as a function of the spectral overlap of competing talkers. *J Acoust Soc Am* (in press)
- BEST V, MARRONE N, MASON CR, KIDD G JR (2012) The influence of non-spatial factors on measures of spatial release from masking. *J Acoust Soc Am* 13:3103–3110
- BOLIA RS, NELSON WT, ERICSON MA, SIMPSON BD (2000) A speech corpus for multitalker communications research. *J Acoust Soc Am* 107:1065–1066
- BRUNGART DS (2001) Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am* 109:1101–1109
- BRUNGART DS, SIMPSON BD, ERICSON MA, SCOTT KR (2001) Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J Acoust Soc Am* 110:2527–2538
- BRUNGART DS, CHANG PS, SIMPSON BD, WANG DL (2006) Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation. *J Acoust Soc Am* 120:4007–4018
- BRUNGART DS, CHANG PS, SIMPSON BD, WANG DL (2009) Multitalker speech perception with ideal time-frequency segregation: effects of voice characteristics and number of talkers. *J Acoust Soc Am* 125:4006–4022
- BYRNE DJ, PARKINSON A, NEWALL P (1990) Hearing aid gain and frequency response requirements for the severely-profoundly hearing impaired. *Ear and Hearing* 11:40–49
- CULLING JF, HAWLEY ML, LITOVSKY RY (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
- DILLON H (2012) Hearing aids, 2nd edn. Boomerang Press, Turramurra
- DUBNO JR, AHLSTROM JB, HORWITZ AR (2008) Binaural advantage for younger and older adults with normal hearing. *J Speech Lang Hear Res* 51:539–556
- FESTEN JM, PLOMP R (1990) Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am* 88:1725–1736
- FREYMAN RL, HELFER KS, BALAKRISHNAN U (2007) Variability and uncertainty in masking by competing speech. *J Acoust Soc Am* 12:1040–1046
- GEORGE ELJ, FESTEN JM, HOUTGAST T (2006) Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 120:2295–2311
- GLYDE H, HICKSON L, CAMERON S, DILLON H (2011) Problems hearing in noise in older adults: a review of spatial processing disorder. *Trends Amplif* 15:116–126

- GLYDE H, CAMERON S, DILLON H, HICKSON L, SEETO M (2013) The effects of hearing impairment and aging on spatial processing. *Ear and Hearing* 34:15–28
- HALL JW, BUSS E, GROSE JH, ROUSH PA (2012) Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear and Hearing* 33:340–348
- HELPER KS, FREYMAN RL (2008) Aging and speech-on-speech masking. *Ear and Hearing* 29:87–98
- KIDD G JR, MASON CR, RICHARDS VM, GALLUN FJ, DURLACH NI (2008) Informational masking. In: Yost WA, Popper AN, Fay RR (eds) *Auditory perception of sound sources*. Springer Handbook of Auditory Research, New York, pp 143–190
- LEVITT H, RABINER LR (1967) Predicting binaural gain in intelligibility and release from masking in speech. *J Acoust Soc Am* 42:820–829
- LICKLIDER JCR (1948) The influence of interaural phase relations upon the masking of speech by white noise. *J Acoust Soc Am* 20:150–159
- MARRONE N, MASON CR, KIDD G JR (2008) 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
- MOORE BCJ (2007) *Cochlear hearing loss: physiological, psychological, and technical issues*, 2nd edn. Wiley, Chichester
- NOBLE W, BYRNE D, TER-HOST K (1997) Auditory localization, detection of spatial separateness, and speech hearing in noise by hearing impaired listeners. *J Acoust Soc Am* 102:2343–2352
- POLLACK I, PICKETT JM (1958) Stereophonic listening and speech intelligibility against voice babble. *J Acoust Soc Am* 30:131–133
- SPENCER NJ, DURLACH NI, COLBURN HS (2012) Relating performance in speech intelligibility tasks with spatially separated maskers to performance in basic binaural tasks. In: *Proc. Assoc. Res. Otolaryng.* San Diego