

Listening in a multisource environment with and without hearing aids

NICOLE L. MARRONE, CHRISTINE R. MASON, AND GERALD KIDD, JR.

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

The aim of the current study was to examine the challenge faced by listeners with hearing loss when selectively attending to one source in the presence of multiple competing sources and reverberation. In a series of experiments, both younger and older listeners with normal hearing or bilateral symmetric sensorineural hearing loss served as subjects. The listeners with hearing loss were experienced users of bilateral hearing aids and were tested unaided, bilaterally aided, and unilaterally aided. The task was to repeat key words spoken by a target talker located straight ahead in the presence of two colocated or symmetrically spatially separated competing talkers. On average, listeners with normal hearing demonstrated a large benefit of spatial separation which was somewhat reduced when the room reverberation was increased. The presence of bilateral sensorineural hearing loss decreased this benefit in both room conditions. Listening through bilateral personal amplification was not significantly different from unaided listening (at an adequate sensation level). However, when listening with one ear aided and one ear unaided the already small benefit was somewhat reduced. Current results suggest an interaction between peripheral hearing loss, hearing aid use, reverberation and performance in an auditory spatial attention task and present a challenge to current models.

INTRODUCTION

The problem of understanding speech in background noise or with other talkers is a common complaint of listeners with hearing loss seeking care in audiology clinics and there is strong evidence that this problem is not fully remediated by amplification (e.g., Gatehouse and Noble, 2004; Harkins and Tucker, 2007). One factor, the spatial separation of sound sources, normally provides a substantial release from masking in normal hearing listeners (see reviews by Yost, 1997, Bronkhorst, 2000, and Ebata, 2003) and thus enhances selective listening. It is also evident that spatial hearing for speech masked by speech is a very different problem from speech masked by noise (e.g., Kidd *et al.*, 2007). However, very little is known about spatial release from speech-on-speech masking in listeners with hearing loss and even less is known about the benefits that unilateral or bilateral hearing aids might provide in this circumstance. The purpose of the present study was to examine spatial release from masking in listeners with normal hearing and listeners with hearing loss in both unaided and aided conditions using speech maskers in realistic conditions such as in the presence of room reverberation.

At the start of the current project consisting of a series of experiments, Marrone *et al.* (2007a) found evidence of spatial tuning and a large spatial release from masking in a

group of young normal hearing listeners that averaged 12 dB when two speech maskers were symmetrically placed at $\pm 90^\circ$ and a similar speech target was directly ahead at 0° azimuth. They concluded that a large component of this spatial release from speech on speech masking was a reduction in informational masking due primarily to the use of interaural differences enabling the listener to perceptually segregate and attend to the target talker while effectively ignoring or filtering out the masker talkers. This effect is in contrast to the benefits provided by lower-level mechanisms such as better-ear listening and within-channel binaural unmasking usually associated with decreases in energetic masking. For instance, in a control condition, a much smaller benefit of spatial separation was found when the speech maskers were replaced by speech-shaped, speech-modulated noise maskers.

In order to determine whether listeners with hearing loss would also show this large benefit of spatial separation when the maskers are other talkers, the approach used in this study (which was nearly identical to that used by Marrone *et al.*, 2007a and is described further in Marrone *et al.*, 2007b) was to create large amounts of informational masking while minimizing certain acoustic factors such as better-ear listening. In this study, we examined spatial release from masking in forty listeners, twenty of whom had bilateral symmetric sensorineural hearing loss and regularly wore two hearing aids. The other twenty listeners were age-matched normal-hearing controls. For the listeners with hearing loss, spatial release from masking was measured unaided, with both aids worn, and with one aid worn. For the normal hearing listeners, spatial release was measured while listening with both ears and with one ear fitted with an earplug and earmuff producing a simulated “monaural” condition.

METHODS

Listeners

There were two main groups of twenty subjects, a group with normal hearing (NH) and a group with hearing loss (HL). The normal hearing listeners had audiometric thresholds ≤ 25 dB HL at octave frequencies from 250-4000 Hz. The listeners with hearing loss had mild-to-moderately severe flat or gradually sloping symmetric sensorineural hearing loss and regularly wore two hearing aids. In both the NH and HL groups there were 10 younger (ages 19-42) and 10 older (ages 57-80) participants. Despite the fact that the older listeners in the NH group met the standards for “clinically normal” hearing, their pure tone audiometric thresholds were on average about 10 dB worse across frequencies than their younger counterparts. In the HL group the younger and older listeners also showed a small difference in the direction of poorer hearing for the older listeners. This point will become important in the discussion of the results below. In the aided conditions the HL listeners were tested with their own aids at their normal settings.

Stimuli

The stimuli were recordings of the four female talkers from the Coordinate Response Measure (CRM) corpus (Bolia *et al.*, 2001) of sentences with the structure, “Ready

[callsign] go to [color] [number] now.” The corpus has eight callsigns, four colors, and eight numbers. On every trial, the listener heard three of the sentences which varied from trial to trial. The target sentence was identifiable by the callsign “Baron”. The talkers, callsigns, colors, and numbers were all mutually exclusive. The task was to report the color-number combination spoken by the target talker (the talker uttering “Baron”).

Description of amplification used and electroacoustic measures

In the aided conditions of the study the HL listeners wore their own hearing aids with their regular fittings in omnidirectional mode. The hearing aids were required to be in working order but no attempt was made to change or improve the fitting. This was based on the decision to have a sample that was representative of the hearing aids/fittings in current use by the listeners recruited for the study. Electroacoustic measurements of each hearing aid were made at each of the two listening sessions. Coupler measurements were made using a Frye Systems 7000 test box and hearing aid analyzer to characterize the frequency response, determine input/output transfer functions, assess attack/release time and processing delay, and to determine gain at user settings. Following otoscopic examination, hearing aid fittings were verified through probe microphone measurements in the listener’s ear using the speech-modulated, speech-weighted noise test signal (“digital speech”) on the Frye Systems 7000 real ear analyzer. These measurements also verified that there was no significant change in the hearing aids or settings across sessions.

Room conditions

The study was conducted in a large single-walled IAC sound booth (12’4” long, 13’ wide, and 7’6” high) that was designed to allow changing the sound absorption characteristics of its surfaces. This is done by covering ceiling, walls, floor, and door with panels of different acoustic reflectivity, such as acoustic foam or Plexiglas®. For the current experiment, the surfaces were left uncovered for one condition (referred to as “BARE”) and for the other room condition all surfaces were covered with reflective panels (“PLEX”). The BARE condition was that of a standard IAC booth: the ceiling, walls, and door had a perforated metal surface and the floor was carpeted. In the PLEX condition, the ceiling, walls, door, and floor were covered with Plexiglas© panels, creating a noticeable increase in reverberation when entering the room and approximately a four-fold increase in reverberation time. These are the same two room conditions found in Kidd *et al.* (2005).

Procedures

The experimental set-up and procedures used were similar to those of Marrone *et al.* (2007a,b). Listeners were seated in the sound booth with an array of 7 loudspeakers arranged in a semicircle in the horizontal plane. The stimuli were presented (in various conditions) from 3 of those loudspeakers including the one positioned in front of the listener (0°) and the two at either side ($\pm 90^\circ$). The loudspeakers were at a distance of 5 feet from the approximate location of the center of the listener’s head when seated.

The equipment used to control the experiment was located outside the booth (computer controlled TDT hardware).

The task was 1-interval closed set (4 x 8 alternative) identification with feedback. Listeners used a handheld keypad with LCD display (Q-term II) to enter their responses and receive feedback on each trial. They were instructed to identify the color and number from the sentence with the callsign “Baron” and were informed that this sentence would always be presented from the loudspeaker directly ahead. Responses were scored as correct only if the listener identified both the color and number accurately. Listeners completed a short practice block of target identification in quiet at a comfortable listening level to familiarize them with the procedures and keypad.

Testing in each room condition began with two quiet conditions. First, unmasked identification thresholds for the target CRM sentences at the target location (0°) were obtained. Second, percent correct performance in an unmasked fixed level identification task was measured for the target at the level at which it would be presented in the masked conditions. There were four masked conditions tested for the HL listeners: binaural unaided, bilateral aided, right ear aided, and left ear aided. For the NH listeners there were three masked conditions, binaural and two with an earplug and earmuff (right ear occluded or left ear occluded). The plug and muff condition was intended as a “monaural” control and a potentially useful comparison for the unilateral aided conditions in the HL group. These conditions were tested in both room conditions (in different order for different listeners) on different days.

For the NH listeners, the target level in all masked conditions was set to 60 dB SPL. For the HL listeners, the target was set to 30 dB sensation level (SL) re. the quiet speech identification thresholds (for CRM sentences) whenever possible and at lower SLs in a few cases (for more on SL see Marrone *et al.*, 2007b). The level of the maskers was then varied adaptively in all masked conditions to estimate threshold (performance of 50% correct identification for both color and number). To facilitate comparisons across listeners and conditions, the thresholds are expressed in target-to-masker ratio in dB where 0 dB T/M indicates that the target was at the same level as each of the individual two masker talkers. This metric is also commonly referred to as a speech recognition threshold or SRT.

In the masked conditions, listeners heard three sentences played concurrently (one target with two independent maskers) on every trial. The maskers were either colocated with the target at 0° or symmetrically spatially separated (target at 0°, one masker from -90° and the other masker from +90°). There was no difference in level between the two masker talkers. Final thresholds were taken as the average of the estimates from four adaptive tracks. There were no obvious training or learning effects seen and thus the results are based on averages of all threshold estimates.

RESULTS

CRM thresholds in quiet for the BARE room condition were highly correlated with both the audiometric pure-tone average (PTA; $r=0.94$, $p<0.001$) and the standard audiometric speech recognition threshold (SRT; $r=0.96$, $p<0.001$) obtained using spondaic words (the PTA and SRT values used were the average of values for the right and left ears). In both room conditions and for all listeners, speech identification performance in quiet was nearly perfect when the CRM sentences were presented at the test level to be used for the target.

Masked results are expressed in terms of T/M and the difference in the T/M at threshold for the colocated and spatially separated conditions is the spatial release from masking (SRM). This series of experiments was initially designed to also examine the effect of age (see Marrone *et al.*, 2007b). As mentioned in the section describing the listeners, each of the two main groups (NH and HL) contained subgroups of younger and older participants. An initial analysis of the unaided data revealed a significant correlation ($r=-0.82$, $p<0.001$) between SRM and “amount of hearing loss” (as estimated by their CRM thresholds in quiet; using their PTA or SRT for spondees measured audiometrically gave similarly high correlations). The slope of this function revealed a relationship such that a 10 dB increase in quiet threshold was associated with a 2dB reduction in SRM. Since the mean difference in SRM between younger and older listeners for both the NH and HL groups was on this order, the effect of age in and of itself could not be reliably assessed. It is possible that these small differences in hearing status alone were responsible for the differences seen across age groups. Therefore, the results presented here are combined across younger and older listeners in both the NH and HL groups.

The group-mean results are contained in Table 1. The entries are T/M at threshold (in dB) for the colocated (0°) and spatially separated ($\pm 90^\circ$) conditions and the associated difference as SRM (in dB). For the unilateral aided conditions as well as the monaural control condition the results were averaged over the estimates for right and left ears.

NH Binaural vs. HL Binaural Unaided

When the three talkers were colocated (0° separation), performance was similar across groups and room conditions. Thresholds were at a target level that was slightly higher than that of the combined maskers (a T/M of 4-6 dB). Although the NH group obtained slightly lower thresholds than the HL group in both rooms, the differences were not statistically significant. The effects of hearing status and room condition were much greater in the spatially separated condition. When listening binaurally, the group mean thresholds for the NH group are 7.8 dB lower than for the HL group in the BARE room condition and 5.1 dB lower in the PLEX room condition. Moreover, thresholds increased for the NH listeners by 4.4 dB in the PLEX room relative to the BARE room with a corresponding increase in threshold of 1.7 dB for the HL group.

		BARE			PLEX		
		T/M 0°	T/M 90°	SRM	T/M 0°	T/M 90°	SRM
NH	Binaural	3.9 (1.0)	-6.1 (3.8)	10.1 (3.3)	4.1 (0.8)	-1.7 (3.3)	5.9 (3.0)
	"Monaural"	3.7 (0.9)	4.3 (1.7)	-0.6 (1.5)	4.4 (1.1)	5.2 (1.6)	-0.8 (1.2)
HL	Binaural Unaided	5.6 (2.0)	1.7 (3.3)	3.9 (2.7)	5.6 (2.2)	3.4 (3.6)	2.2 (2.1)
	Bilateral Aided	5.1 (1.4)	2.1 (3.9)	3.1 (2.9)	5.6 (1.9)	3.9 (4.2)	1.6 (2.6)
	Unilateral Aided	4.9 (1.4)	2.7 (3.8)	2.2 (2.9)	5.5 (2.0)	4.9 (3.4)	0.6 (1.9)

Table 1: Group mean results (bold) and one standard deviation (parentheses).

A comparison of the results for the colocated condition with those for the spatially separated condition can be summarized in terms of the amount of SRM, a measure of the benefit of spatial separation. Fig. 1 shows group mean SRM in dB for listener group, listening condition and room condition. The error bars are one standard error of the mean over listeners. Both the NH (listening binaurally) and HL (unaided) groups showed a benefit of spatial separation between the target and competing talkers in both room conditions. However, the amount of benefit was dependent upon both listener group and room condition with the NH group obtaining much greater SRM than the HL group in both room conditions (first two bars in each group of bars) and the added reverberation in the PLEX room decreasing the amount of SRM for both listener groups. More details about the results for the NH and HL groups can be found in Marrone *et al.*, 2007b. The results of the monaural control condition (shown in both the table and figure), in which no SRM was obtained, indicate that the phenomenon depends on binaural hearing, i.e., monaural cues are not sufficient to do the task.

Unaided vs. Bilateral aided and Unilateral aided

When the three talkers were colocated, the average T/Ms at threshold were again remarkably similar across listening condition and room condition. However, a slightly larger range of performance was observed when the three talkers were spatially separated. When the maskers were located at $\pm 90^\circ$, the lowest (best) T/Ms at threshold occurred in the BARE room for the unaided condition whereas the highest (worst) T/Ms at threshold occurred in the unilateral aided conditions in the PLEX room. In all cases, both aided and unaided, the HL listeners required positive T/Ms (target talker higher than masker talkers) to achieve 50% correct.

For a given amount of room reverberation, the amount of benefit from spatial separation is not significantly different when listening with two hearing aids than when listening without aids at a similar sensation level. Despite the small differences, the SRM

for listening unilaterally is significantly reduced compared to listening binaurally (unaided) or bilaterally aided. In looking at individuals, there was a moderately strong correlation between the amount of SRM in the unaided condition and the amount of SRM in the bilateral aided condition in both the BARE room ($r=0.74, p<0.001$) and in the PLEX room ($r=0.68, p<0.001$). In addition, the correlation between unilateral and bilateral aided results was also significant ($r = 0.74, p<0.001$ in BARE and $r=0.67, p<0.001$ in PLEX).

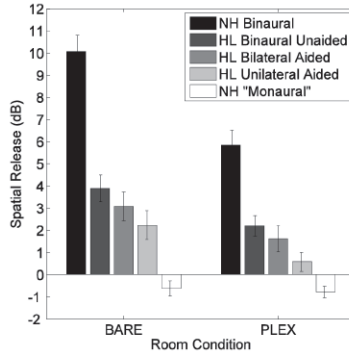


Fig. 1: Group mean spatial release from masking in dB. Error bars are plus and minus one standard error of the mean across listeners.

CONCLUSIONS AND DISCUSSION

Consistent with the report by Marrone *et al.* (2007a), NH listeners demonstrated a large benefit of spatial separation (10dB) when selectively listening to one of multiple simultaneous talkers (see also Marrone *et al.*, 2007b). This effect was obtained without the availability of a simple acoustic better ear advantage, as indicated by the absence of SRM in the simulated monaural condition. The effect was also somewhat robust with respect to increased reverberation in that a substantial SRM was still observed (6dB) in the more reverberant condition. Bilateral sensorineural hearing loss significantly decreased the benefit of spatial separation. Both listener groups had similar T/Ms at threshold when the three talkers were colocated. However when the talkers were spatially separated, listeners with hearing loss required substantially higher T/Ms at threshold than normal-hearing listeners in both room conditions. On average, listeners with hearing loss required the target talker to be higher in level than the other two talkers to achieve 50% correct whereas the NH listeners could achieve this performance when the masker talkers were higher in level than the target. For some listeners with hearing loss, performance was as poor in the spatially separated condition as in the colocated condition. One possible explanation for the much reduced SRM in HL listeners is increased energetic masking due to broader auditory filters. Evidence of this in listeners with sensorineural hearing loss with a concomitant reduction in SRM has been reported by Arbogast *et al.* (2005). In that study however, there was a single masker talker positioned on one side of the listener allowing for the contribution from acoustic head shadow, and the target and maskers were in different frequency bands.

If all or part of the explanation is in EM differences then it is not completely clear why there is a differential effect for the colocated and spatially separated conditions. One explanation is that the results for the colocated condition are already at some sort of ceiling; that is, once the target talker is sufficiently higher in level than the other two talkers, the level alone is a sufficient segregation cue in all listener groups, listening conditions and room conditions. It is also difficult to predict how large a difference in EM might realistically be expected for these stimuli and what consequence it would have on performance. Future studies, now underway and including more measurements with modulated and unmodulated noise maskers, as well as noisy speech maskers, will attempt to isolate these components to a greater degree. Another (related) possibility is that the presumed degraded spectral and temporal representation of the speech as a consequence of the sensorineural hearing loss affects the ability of the HL listeners to segregate the target stream and maintain it over time (during the course of the sentence) and/or direct attention to it. This is in accordance with the view that segregation is not an all or none phenomenon but rather varies in strength according to the clarity of the representation of the sources and the usefulness of the segregation cues as well as the assumption that attention is given to well formed objects (for discussion see Shinn-Cunningham, this volume). In the spatially separated case this implies a reduced ability to use interaural differences to segregate and maintain the target speech stream. This general view is also consistent with the findings here of reduced SRM with increased reverberation. Overall, these results illustrate one reason why listeners with hearing loss often report such great difficulty understanding a target talker in real rooms when there are other interfering talkers. Unlike listeners with normal hearing, they often receive little benefit from spatial separation of sound sources. This is presumably because they are unable to effectively focus on the target to the exclusion of the interferers (for whatever reason, including the fact that there is more residual EM that cannot be undone by binaural cues, implying that the target and maskers are “mixed” to a greater degree for these listeners).

On average, the amount of spatial release from masking with bilateral hearing aids was not significantly different than unaided performance at a comparable sensation level. Although it was possible that bilateral amplification could improve spatial release from masking, for instance by restoring more high frequency information, this was not the case. Real ear measurements with and without their hearing aids were analyzed and it was determined that most subjects realized substantially more (frequency specific) gain in the aided conditions, however this did not translate into a performance benefit for these conditions. It is possible that listening through bilateral aids could have had the opposite effect if the two aids distorted important interaural timing or level differences, or were poorly matched such that integration of information across the ears was adversely affected. The current result should not be interpreted as implying that the hearing aid does not provide benefit in this type of listening situation since the presentation level in the unaided condition was at a comparable SL and is therefore not representative of unaided listening in daily life in which much information could be inaudible. This result is promising, given that performance with hearing aids could have been worse than without, as has been observed in studies of

horizontal localization (e.g. Van den Bogaert *et al.*, 2006). This suggests that there is room for improvement in aided speech recognition in the presence of multiple competing talkers. This would only be the case, however, if the fundamental limitation on spatial release from masking is not determined by the degree and nature of the hearing impairment. It may be that the best that can occur with hearing aid fittings is to preserve whatever spatial release from masking is present unaided unless the amplification provided can strengthen perceptual segregation cues even in the face of potentially greater amounts of energetic masking.

The amount of spatial release from masking obtained with an unilateral aid, which was slightly but significantly poorer than with bilateral aids or unaided at an equivalent SL, was related to the amount of spatial release obtained with bilateral hearing aids. Given the “monaural” result with the NH listeners it was initially surprising that any spatial release could be obtained with one hearing aid. However, it became clear that the sensation level of the target in the unaided ear was higher for most HL listeners than it was for the NH listeners wearing an earplug and earmuff. The listeners with milder hearing loss were apparently still able to make use of binaural information and achieve some, albeit reduced, spatial release from masking in the unilateral aided listening condition. There was an interaction between aided listening condition and the amount of room reverberation such that the best performance when using hearing aids was obtained for a bilateral fitting in low reverberation and the worst performance occurred when listening with a single hearing aid in higher reverberation.

ACKNOWLEDGMENTS

This work was supported by grant number FA9950-05-1-2005 from the Air Force Office of Scientific Research and by grants DC00100, DC04545 and DC04663 from the National Institute on Deafness and other Communication Disorders (NIDCD). The first author was supported by grant number F31 DC008440 from NIDCD. This work partly fulfills the requirements for the doctor of philosophy degree at Boston University by the first author. Portions of this work were presented at the meeting of the American Auditory Society in Scottsdale, AZ, USA in March, 2007 and the meeting of the Acoustical Society of America in Salt Lake City, UT, USA in June, 2007. The authors wish to thank the listeners that participated in the experiment for their time and effort, as well as our colleagues at BU, particularly Virginia Best, H. Steven Colburn, Nathaniel Durlach, Melanie Matthies, and Barbara Shinn-Cunningham for many helpful discussions. Finally, Deborah Corliss, Janine Francis and Jacquelyn Therieau Stachel are thanked for their assistance with the Plexiglas® in the Sound Field Lab.

REFERENCES

- Bolia, R. S., Nelson, W. T., Ericson, M. A., and Simpson, B. D. (2000). “A speech corpus for multitalker communications research,” *J. Acoust. Soc. Am.* **107**, 1065-1066.
- Bronkhorst, A. (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acust. Acta Acust.*, **86**, 117-128.

- Ebata, M. (2003). "Spatial unmasking and attention related to the cocktail party problem," *Acoust. Sci. Tech.* **24**, 208-219.
- Gatehouse, S., and Noble, W. (2004). "The Speech, Spatial, and Qualities of Hearing Scale (SSQ)," *Int. J. Audiol.* **43**, 85-99.
- Harkins, J., and Tucker, P. (2007). "An internet survey of individuals with hearing loss regarding assistive listening devices," *Trends Ampl.*, **11**, 91-100.
- Kidd, J. G., Mason, C. R., Brughera, A., and Hartmann, W. M. (2005). "The role of reverberation in release from masking due to spatial separation of sources for speech identification," *Acust. Acta Acust.*, **91**, 526-536.
- Kidd, G. Jr., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. (2007). "Informational masking," in *Auditory Perception of Sound Sources*, edited by W.A. Yost (Springer, New York).
- Marrone, N. L., Mason, C. R., and Kidd, G. Jr. (2007a). "Tuning in the spatial dimension: Evidence from a masked speech identification task," (under review).
- Marrone, N. L., Mason, C. R., and Kidd, G. Jr. (2007b). "Effect of hearing loss on the benefit of spatial separation between multiple talkers in reverberant rooms," (under review).
- Van den Bogaert, T., Klasen, T. J., Moonen, M., Van Deun, L., and Wouters, J. (2006). "Horizontal localization with bilateral hearing aids: without is better than with," *J. Acoust. Soc. Am.* **119**, 515-526.
- Yost, W. A. (1997). "The cocktail party problem: Forty years later," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. A. Gilkey and T. R. Anderson. Mahwah, NJ, Lawrence Erlbaum, 329-348.