

# The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms

Nicole Marrone, Christine R. Mason, and Gerald Kidd

Citation: [The Journal of the Acoustical Society of America](#) **124**, 3064 (2008); doi: 10.1121/1.2980441

View online: <https://doi.org/10.1121/1.2980441>

View Table of Contents: <https://asa.scitation.org/toc/jas/124/5>

Published by the [Acoustical Society of America](#)

---

## ARTICLES YOU MAY BE INTERESTED IN

[The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners](#)

The Journal of the Acoustical Society of America **117**, 2169 (2005); <https://doi.org/10.1121/1.1861598>

[The effect of spatial separation on informational and energetic masking of speech](#)

The Journal of the Acoustical Society of America **112**, 2086 (2002); <https://doi.org/10.1121/1.1510141>

[Tuning in the spatial dimension: Evidence from a masked speech identification task](#)

The Journal of the Acoustical Society of America **124**, 1146 (2008); <https://doi.org/10.1121/1.2945710>

[The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer](#)

The Journal of the Acoustical Society of America **115**, 833 (2004); <https://doi.org/10.1121/1.1639908>

[Spatial release from informational masking in speech recognition](#)

The Journal of the Acoustical Society of America **109**, 2112 (2001); <https://doi.org/10.1121/1.1354984>

[The influence of non-spatial factors on measures of spatial release from masking](#)

The Journal of the Acoustical Society of America **131**, 3103 (2012); <https://doi.org/10.1121/1.3693656>

---



The banner features a blue background with a close-up image of a 3D printer nozzle depositing material into a red, lattice-like structure. On the left, the JASA logo is displayed in white, with the text 'THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA' underneath. To the right of the logo, the text 'Special Issue:' is in yellow, and 'Additive Manufacturing and Acoustics' is in white. At the bottom left, a yellow button contains the text 'Submit Today!'.

**JASA**  
THE JOURNAL OF THE  
ACOUSTICAL SOCIETY OF AMERICA

**Special Issue:**  
**Additive Manufacturing and Acoustics**

[Submit Today!](#)

# The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms

Nicole Marrone,<sup>a)</sup> Christine R. Mason, and Gerald Kidd, Jr.

*Department of Speech, Language, and Hearing Sciences and the Hearing Research Center, Boston University, 635 Commonwealth Avenue, Boston, Massachusetts 02215*

(Received 3 August 2007; revised 30 July 2008; accepted 31 July 2008)

This study investigated the interaction between hearing loss, reverberation, and age on the benefit of spatially separating multiple masking talkers from a target talker. Four listener groups were tested based on hearing status and age. On every trial listeners heard three different sentences spoken simultaneously by different female talkers. Listeners reported keywords from the target sentence, which was presented at a fixed and known location. Maskers were colocated with the target or presented from spatially separated and symmetrically placed loudspeakers, creating a situation with no simple “better-ear.” Reverberation was also varied. The target-to-masker ratio at threshold for identification of the fixed-level target was measured by adapting the level of the maskers. On average, listeners with hearing loss showed less spatial release from masking than normal-hearing listeners. Age was a significant factor although small differences in hearing sensitivity across age groups may have contributed to this effect. Spatial release was reduced in the more reverberant room condition but in most cases a significant advantage remained. These results provide evidence for a large benefit of spatial separation in a multitalker situation that is likely due to perceptual factors. However, this benefit is significantly reduced by both hearing loss and reverberation.

© 2008 Acoustical Society of America. [DOI: 10.1121/1.2980441]

PACS number(s): 43.66.Dc, 43.66.Lj, 43.66.Pn, 43.66.Sr [RLF]

Pages: 3064–3075

## I. INTRODUCTION

It is well documented that spatial separation of one or more interfering sound sources from a target source may result in significant release from masking in normal-hearing listeners [see reviews by Yost (1997), Bronkhorst (2000), Ebata (2003), and Colburn *et al.* (2006)]. In general, there have been fewer psychophysical studies of spatial release from masking in listeners with hearing loss, especially when the interfering sources are competing talkers in real rooms. However, that type of situation is known to be among the more challenging communication environments for listeners with hearing loss and as such it is important to fully understand the factors underlying their difficulties. Gatehouse and Noble (2004) designed the speech, spatial, and qualities of hearing scale (SSQ) to assess the perceived difficulty of listening in complex and dynamic auditory environments. They found that the items related to aspects of auditory selective attention and switching attention (e.g., having a conversation with one person when many people are talking, following the TV while talking to someone, etc.) were those most highly correlated with experiencing handicap as a result of hearing loss. Thus, it would seem that there is a compelling rationale for studying selective listening in listeners with hearing loss in complex sound environments.

Traditionally, spatial release from masking has been thought to be a consequence of relatively low-level (i.e., peripheral) factors, including the simple acoustic effect of “better-ear listening” [attending to the ear with the higher

target-to-masker (T/M) ratio] and binaural analysis (leading to “bottom-up” within-channel improvements in T/M, or “masking level differences”). Models of this process based on interaural level differences (ILDs) and interaural time differences (ITDs) are generally successful in predicting the improvement in the intelligibility of target speech masked by noise when the two sources are spatially separated (e.g., Zurek, 1993; Bronkhorst, 2000). Yet when the listening situation is more complex and the maskers are highly distracting similar sounds such as one or more competing streams of speech (often producing large amounts of centrally based “informational masking,” cf. Kidd *et al.*, 2008), a large release from masking can be observed, which is not accounted for by lower-level peripheral factors and is not successfully predicted by such models. The results of several recent studies using more complex stimuli and challenging listening environments have shown that spatial release from masking is a complicated phenomenon comprising both lower-level and higher-level processes (e.g., Kidd *et al.*, 1998; Freyman *et al.*, 1999; Hawley *et al.*, 1999; Noble and Perrett, 2002; Arbogast *et al.*, 2002; Hawley *et al.*, 2004; Best *et al.*, 2006; Marrone *et al.*, 2008). Currently, very little is known about the role of nonperipheral factors in complex listening situations in listeners with hearing loss. Thus, the goal of the current study was to examine higher-level factors (e.g., selective attention) in spatial release from masking and determine the extent to which the benefit of spatially separating competing talkers is affected by sensorineural hearing loss and age.

The use of a single masking source or the asymmetric positioning of multiple masker sources to create spatial release may make the separation of the contributions of lower-

<sup>a)</sup>Author to whom correspondence should be addressed. Electronic mail: n-marrone@northwestern.edu

level and higher-level processes difficult to ascertain because of the resulting opportunity for the listener to benefit from the improvement in T/M in one ear due to acoustic head shadow. Consequently, the maskers in the current study were placed symmetrically around the target in order to diminish or eliminate any potential for a better-ear advantage. Furthermore, the stimuli and methods used are known to produce large amounts of informational masking, which in young normal-hearing listeners may be substantially reduced by spatially separating the sound sources (Arbogast *et al.*, 2002; Brungart and Simpson, 2002; Shinn-Cunningham *et al.*, 2005; Marrone *et al.*, 2008). The main finding of the study of Marrone *et al.* (2008), which employed the same stimuli and similar methods to those used here, was a tuned response in azimuth. More specifically, the T/M for speech identification threshold progressively decreased as the two symmetric speech maskers were varied in azimuth from 0° (colocated with target) to 45°. The benefit of spatial separation between the target and maskers had an asymptotic maximum of 12 dB on average. Importantly, a control condition in which listeners had one ear occluded by an earplug and earmuff revealed that there was no release from masking with increasing spatial separation of the maskers when the listener could only listen monaurally. Although binaural cues are thus fundamentally important to the task, within-channel binaural analysis was not considered to be a significant factor in producing the spatial release observed in that study because it was concluded that (a) relatively little within-channel energetic masking was present (cf. Brungart *et al.*, 2006), suggesting that masking level differences (MLDs) would be minimal; (b) the amount of release predicted by the Bronkhorst (2000) model (for speech in noise, approximately 2 dB) was much smaller than that which was observed, also consistent with the notion that the task was dominated by informational masking; and (c) a noise control condition (producing primarily energetic masking) resulted in very little spatial release (about 1–2 dB), in agreement with that model's prediction.

Marrone *et al.* (2008) speculated that segregating the three talkers was a very difficult task when the talkers were colocated primarily because of the high degree of target-masker similarity (all talkers were female) as well as target talker uncertainty (the target talker was chosen randomly from among them on each trial). In this difficult listening situation, identification of the target speech was only reliably possible when the level of the target was higher than that of the maskers. In support of that idea, they simply reversed the masker speech, thereby decreasing the target-masker similarity (and the opportunity for target-masker confusion), and improved the threshold T/M significantly (by about 12 dB). Because of the possibility that energetic masking could have differed for forward and reversed speech maskers (due to overall envelope differences or because the target and masker words are aligned differently in the two conditions), Marrone *et al.* (2008) also tested a control condition using speech-shaped speech-envelope-modulated noise as a masker. The speech envelope was taken from either forward or reversed speech maskers and the amount of masking produced in the two conditions was compared. The results of

that control condition supported the conclusion that the difference for forward versus backward speech maskers was not due to differences in energetic masking. Furthermore, a release from masking for the forward speech of nearly 8 dB was observed despite presentation in a more reverberant listening condition. For the reasons listed above and others considered in that article, Marrone *et al.* (2008) concluded that the spatial release observed was due primarily to improved speech stream segregation and stronger focus of attention along the spatial dimension. Although this process is still not fully understood and clearly depends on interaural differences, Marrone *et al.* (2008) hypothesized that the presence of a high degree of informational masking was necessary (although perhaps not sufficient) to observe large effects with symmetric speech maskers.

Several previous studies have examined the utility of spatial cues in overcoming masking in listeners with hearing loss. Generally, these studies have used noise or multitalker babble as maskers and have found smaller spatial release in listeners with hearing loss than with normal hearing (Duquesnoy, 1983; Gelfand *et al.*, 1988; Bronkhorst and Plomp, 1989, 1992; Helfer, 1992; Ter-Horst *et al.*, 1993; Peissig and Kollmeier, 1997; Dubno *et al.*, 2002). Noise maskers and multitalker babble may be dominated by energetic masking (cf. Freyman *et al.*, 2004) and so the extent to which these studies are informative about higher-level selective listening processes is unclear. In fact, one frequently cited hypothesis regarding the reduced spatial release from masking in listeners with hearing loss has to do with the reduced opportunity to take advantage of an acoustically better ear. Dubno *et al.* (2002) found that the benefit of spatially separating speech-shaped noise from a target talker was differentially affected by low-pass and high-pass filtering in younger and older adults with normal hearing and older adults with hearing loss. Whereas the two normal-hearing groups demonstrated significant benefit from source separation in both filter conditions, the older adults with hearing loss received little benefit when the speech and noise were high-pass filtered (presumably providing primarily ILD cues) but showed a small benefit from the low-pass speech and noise (probably related primarily to ITDs). This leads to the question of whether the benefit from spatial cues would be similar for listeners with normal hearing and those with hearing loss if no overall better-ear advantage (in contrast to time-varying or momentary, discussed more fully below) were available in the stimulus.

A few of the studies of spatial release from masking in listeners with hearing loss have included competing speech as a masker (Duquesnoy, 1983; Peissig and Kollmeier, 1997; Arbogast *et al.*, 2005; Helfer and Freyman, 2008). These studies have also mostly used single source/asymmetric masker placements and generally have found a decreased benefit of spatial separation in listeners with hearing loss. For example, Duquesnoy (1983) found that the benefit of spatially separating a female talker from a male competing talker by 90° was 6.7 dB on average for young normal-hearing listeners while it was 4.5 dB for older listeners with hearing loss. However, it was not clear whether this difference was due to hearing loss, the difference in age between

the listener groups, or an interaction between these two factors. Furthermore, the relatively small spatial release may have indicated that the different-sex talkers were fairly easy to segregate and produced little informational masking, so the further benefit of providing spatial cues was limited. Peissig and Kollmeier (1997) also tested speech reception in the presence of multiple speech maskers but under conditions of relatively low uncertainty, using known target sentences with continuous interference in a subjective intelligibility procedure to estimate speech threshold with three masking sources. The maximum benefit of spatial separation in this condition was 5 dB on average for the listeners with normal hearing while it ranged from 0.6 to 3.8 dB for the listeners with hearing loss. Arbogast *et al.* (2005) used a form of cochlear implant simulation processing (e.g., Shannon *et al.*, 1995) to filter the target and masker speech into different frequency bands in order to limit energetic masking. They found that the mean spatial release from masking for a 90° horizontal separation between two male talkers was about 15 dB for the listeners with normal hearing and about 10 dB for age-matched listeners with hearing loss. These results were interpreted as evidence that both types of listeners can benefit from spatial separation between a target talker and a speech masker. However, the amount of spatial release was less in the listeners with hearing loss, possibly because of greater energetic masking due to wider auditory filters. Evidence in support of this idea came from an off-frequency noise masking control condition in which the listeners with hearing loss showed significantly greater masking than those with normal hearing. Thus, the asymmetric single masker placement allowed for better-ear cues and the presumably poorer frequency selectivity in the impaired ears may have increased the overlap of excitation between the sharply filtered bands increasing the ratio of energetic to informational masking.

In most of the studies above in which normal-hearing and hearing-impaired listeners were compared, the hearing-impaired listeners were older than their normal-hearing counterparts. There have been a number of studies suggesting that age is a factor in either selective listening tasks (e.g., Helfer and Freyman, 2008) or divided listening tasks (e.g., Humes *et al.*, 2006) separate from the consequences of hearing loss. In a recent study by Dubno *et al.* (2008) it was concluded that older listeners may have deficits in binaural processing for sentence recognition (measured in noise but without better-ear cues by use of symmetric masker placement) even though no age-related differences were found for the tasks involving simple detection and recognition. Furthermore, there are indications that advanced age affects other higher-level processes in hearing such as auditory memory (e.g., Pichora-Fuller *et al.*, 1995). Because the tasks and listening situations we are using in these experiments are explicitly intended to tap higher-level processes such as selective attention and memory, age is an important factor that is incorporated into the current experimental design. In addition, because the majority of available listeners with hearing loss tend to be older, it is important to control for the factor of age separately. Thus, we are attempting to determine

whether older listeners—either with or without hearing loss—exhibit greater difficulty in a complex multitalker environment than matched younger listeners.

The current study applies the methods and stimuli used by Marrone *et al.* (2008) to examine the benefit of spatial separation between multiple competing talkers in younger and older listeners with and without sensorineural hearing loss. The goal is to emphasize the contributions of higher-level processes in spatial release from masking in part through the use of a procedure that severely restricts better-ear acoustic cues and that normally causes large amounts of informational masking. Furthermore, in an attempt to approximate more realistic listening environments where listeners with hearing loss report experiencing the greatest difficulty, masked speech identification is tested under two levels of reverberation.

## II. METHODS

### A. Listeners

Four listener groups were recruited based on hearing status (normal hearing/hearing loss) and age (younger/older). All listeners had symmetric hearing thresholds ( $\leq 10$  dB difference between ears at any audiometric test frequency) and no significant air-bone gap ( $< 10$  dB difference between air and bone conduction thresholds at any audiometric test frequency in one ear). The normal-hearing groups had audiometric thresholds  $\leq 25$  dB hearing level (HL) at octave frequencies from 250 to 4000 Hz. The listeners with hearing loss had mild-to-moderately severe flat or sloping sensorineural hearing loss. All of the listeners with hearing loss wore aids bilaterally but were tested unaided in this experiment. All of the listeners spoke American English as their primary language and were paid for their participation.

The group of ten younger listeners with hearing loss (YHL) had an age range of 19–42 years with a mean age of 28.5 years. They were age-matched with ten younger listeners with normal hearing (YNH), age 22–43 years with a mean age of 31.4 years. The older hearing loss group (OHL) had ten listeners ranging in age from 57 to 80 years with a mean age of 69.9 years. The older normal-hearing (ONH) group had a slightly smaller age range, 59–74 years, but the mean age (66.2 years) was similar to the OHL group. Figure 1 displays the mean audiograms for the four listener groups. Mean thresholds in dB HL across audiometric frequencies from 250 to 8000 Hz are shown for each listener group for both ears. The lowest thresholds across frequency were for the YNH group. The ONH group had slightly higher thresholds at all test frequencies. For the listeners with hearing loss, mean thresholds for the OHL group were within 10 dB HL of those for the YHL group from 250 to 4000 Hz.

### B. Stimuli

The stimuli were recordings of the four female talkers from the coordinate response measure (CRM) corpus (Bolia *et al.*, 2000), in which every sentence has the following structure: “Ready [callsign] go to [color] [number] now.” The target sentence always had the callsign “Baron,” although the talker varied from trial to trial. The maskers were



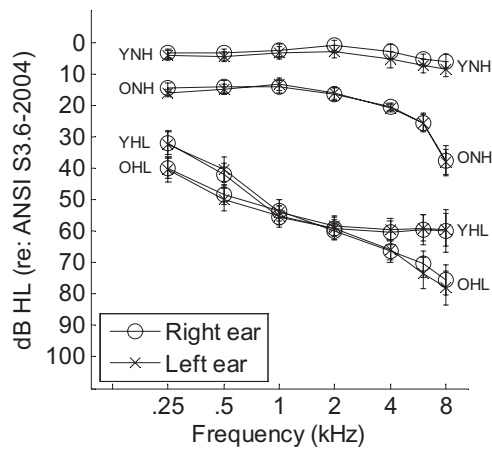


FIG. 1. Mean thresholds in dB HL for each listener group for right and left ears (circle and cross symbols, respectively) plotted as a function of audiometric frequency from 250 to 8000 Hz. The error bars represent  $\pm 1$  standard error of the mean.

also CRM sentences spoken by two different talkers from the corpus. The masker talkers, callsigns, colors, and numbers were different from the target and from each other and varied from trial to trial.

### C. Room conditions

The experiment was conducted in the Soundfield Laboratory at Boston University with stimuli presented through loudspeakers (Acoustic Research 215 PS). The room is a large single-walled Industrial Acoustics Company (IAC) sound booth (12 ft, 4 in long; 13 ft wide; and 7 ft, 6 in. high) in which the reverberation characteristics can be changed by covering all surfaces (ceiling, floor, walls, and door) with custom fit panels of different acoustic reflectivities. All experimental conditions were completed in two sessions, one for each of two room conditions. Session order was counter-balanced across listeners. In the low-reverberation condition, the surfaces were left uncovered (referred to as the “BARE” room condition). This room configuration was that of standard IAC booths: The ceiling and walls had a perforated metal surface and the floor was carpeted. In the second, more reverberant condition, all surfaces were covered with Plexiglas® panels (the “PLEX” room condition). Acoustic measurements of these room configurations were reported by Kidd *et al.* (2005), and included impulse responses, modulation transfer functions, cross-correlation functions, and ILD measurements. There are differences between the two room conditions for each of these measurements corresponding to the increase in reverberation due to the reflective panels in the PLEX room. For example, there was a decrease in the direct-to-reverberant energy ratio of more than 7 dB (from 6.3 to  $-0.9$  dB for BARE versus PLEX, respectively) as measured at the approximate position of the listener’s head. Also, reverberation times increased by a factor of 4 from around 0.06 s in BARE to just over 0.25 s in PLEX.<sup>1</sup>

## D. Procedures

### 1. General

The experimental setup and procedures used were similar to those in Marrone *et al.*, (2008). As mentioned above, there were two sessions each lasting approximately 2 h that were typically scheduled within a few days of each other. Listeners were seated in the sound booth with an array of seven loudspeakers arranged in a semicircle in the horizontal plane; however, stimuli were only presented from the three loudspeakers positioned in front of the listener ( $0^\circ$ ) and to either side ( $\pm 90^\circ$ ) depending on the condition. The faces of the loudspeakers were 5 ft from the approximate location of the center of the listener’s head when seated. Listeners used a handheld keypad with a liquid crystal display (Q-term II) to enter their responses and receive feedback on each trial. The word “listen” appeared on the display at the beginning of each trial. After stimulus presentation, listeners responded to the prompts, “Color [B R W G]?” and “Number [1–8]?” on the keypad. The feedback consisted of a message indicating whether the color and number reported were correct and what the target color and number had been for that trial (for example, “Correct, it was blue four”).

The computer and Tucker-Davis Technologies hardware used to control the experiment and present the stimuli were located outside the booth. Stimuli were played at a 40 kHz rate via a 16 bit, eight-channel digital-to-analog converter, low-pass filtered at 20 kHz, and attenuated. The target was routed through a programmable switch. On trials when the maskers were spatially separated from the target ( $\pm 90^\circ$  spatial separation), each sentence was routed through separate digital-to-analog converter channels, filters, attenuators, power amplifiers (Tascam), and played through separate loudspeakers. On trials when the target and maskers were colocated ( $0^\circ$  spatial separation), the two masker sentences were first digitally added. The combined two-talker masker and the target were then routed through separate digital-to-analog converter channels, filters, and attenuators before being combined in a mixer, passed through a power amplifier, and sent to the loudspeaker.

Prior to each session, the system was calibrated so that the loudspeakers were correctly positioned and the output level measured with a Brüel & Kjær microphone suspended in that position for a given input was verified and the same from each loudspeaker. For a broadband noise of the same level at the input to the loudspeakers measured at the position of the listener’s head, the overall sound pressure level (SPL) was approximately 3 dB higher in the PLEX room than in the BARE room. This correction was made when the results are reported in dB SPL.

The task was a one-interval  $4 \times 8$ -alternative forced-choice with feedback (four colors: red, white, blue, and green and the numbers 1–8). Listeners were instructed to identify the color and number from the sentence with the callsign “Baron” and were informed that this target sentence would always be presented from the loudspeaker directly ahead. Since the target would always be directly ahead of them, they were told to keep their head facing forward, but they were not restrained. 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.

## 2. Preliminary measurements

At the start of each session (for each room condition), a one-up, one-down adaptive procedure was used to estimate the 50% correct point on the psychometric function (Levitt, 1971) for speech identification in quiet (target only, no maskers). In these trials, the listener heard a single CRM sentence with the callsign “Baron” played from the loudspeaker at 0°. The initial step size was 4 dB and after four reversals the step size was decreased to 2 dB. Each adaptive track had a minimum of 30 trials and at least 9 reversals (typically many more than 9 were obtained). The threshold calculation was based on all reversals after the first three, discarding the fourth if necessary to make an even number. Two estimates of threshold in quiet were measured and averaged. If the threshold estimates differed by more than 5 dB an additional two estimates were collected and all four were averaged.

Following threshold measurements in quiet, percent correct speech identification was measured at a fixed target level (target only, no maskers). The groups with normal hearing were tested at 60 dB SPL. The groups with hearing loss were tested, if possible, at 30 dB sensation level (dB SL) with regard to their speech threshold in quiet. The upper limit for stimulus presentation was set at 84 dB SPL. With this in mind, the target level chosen for an individual listener with hearing loss was set high enough to be clearly understood, but low enough to allow the maskers to adapt over a sufficient range for estimating threshold. Some listeners with high quiet thresholds were tested at a lower SL to ensure that the upper limit for presentation was not exceeded by a high-level masker. In cases where the initial level chosen did not yield satisfactory identification in quiet (i.e.,  $\geq 87\%$  correct), the level was increased slightly and fixed-level identification was repeated. For the BARE room conditions, four listeners were tested at 10 dB SL, two were tested at 15 dB SL, five were tested at 20 dB SL, and nine were tested at 30 dB SL. One of the listeners with hearing loss tested at 10 dB SL had slightly lower fixed-level identification (83%) than the other listeners; however, the level could not be increased further due to the equipment limits. Listeners were tested at the same sensation level in the PLEX room as in the BARE room with three exceptions. Two listeners with hearing loss were tested at slightly higher sensation levels in the PLEX room in order to achieve better fixed-level identification scores and one listener was tested at a lower sensation level in PLEX due to comfort. All listeners reported that the target levels used in the main experiment were at comfortable listening levels in both room conditions.

## 3. Speech thresholds with two competing talkers

The conditions comprising the main experiment were tested after the target level was chosen and confirmed to yield highly accurate identification. In these conditions, the

listener heard three sources played concurrently (one target with two independent maskers) on every trial. The trials were blocked by masker location and presented in random order, where masker location changed after every two adaptive tracks. The maskers were either colocated with the target or symmetrically spatially separated (positioned at  $\pm 90^\circ$ ). The target level was fixed at the specified sensation level. The masker level was varied adaptively in 4 dB steps initially and then in 2 dB steps following the fourth reversal. The two masker talkers always had the same rms level. At the beginning of each adaptive track, the target was clearly audible above the maskers (the initial masker level was 20 dB below the target). As in the measurement of quiet threshold, each track had a minimum of 30 trials and at least 9 reversals and the threshold calculation was based on all reversals after the first three, discarding the fourth if necessary to make an even number. Threshold estimates were averaged over four tracks per condition.

## III. RESULTS

### A. Speech identification thresholds in quiet

The speech identification threshold measurements using the CRM stimuli that were made in both room conditions in quiet (without interfering talkers) were obtained for the purpose of setting the target level and are used in later analyses as an approximate estimate of the listener's amount of hearing loss (see Fig. 5). The older groups tended to have slightly higher thresholds in quiet than the younger groups. In both room conditions, threshold in quiet was on average 8 dB SPL higher in the ONH group than in the YNH group and 10 dB SPL higher in the OHL group than in the YHL group.

### B. T/M at threshold with two competing talkers

To facilitate comparisons across conditions, masked thresholds were expressed as T/M in decibels. The T/M was calculated by subtracting the level of the individual maskers at threshold from the fixed target level. Group mean thresholds (and standard error of the mean) for target identification in the presence of two competing talkers for the two spatial separation conditions are shown in Fig. 2. The left panel contains results from the BARE room condition and the right panel shows results from the PLEX room condition. These data were analyzed using a repeated-measures analysis of variance (ANOVA) with two within-subjects factors (spatial separation and room condition) and two between-subjects factors (hearing status and age). Significant main effects were found for all four factors [spatial:  $F(1,36)=236.9$ ,  $p < 0.001$ ; room:  $F(1,36)=43.5$ ,  $p < 0.001$ ; hearing loss:  $F(1,36)=36.9$ ,  $p < 0.001$ ; and age:  $F(1,36)=14.5$ ,  $p = 0.001$ ]. As shown in Fig. 2, the T/Ms at threshold were significantly lower (better) for (a) spatially separated versus colocated talkers, (b) the less reverberant room (BARE), (c) the listeners with normal hearing (open symbols), and (d) the younger listeners within each hearing status category (circles).

A comparison of data in the left and right panels of Fig. 2 illustrates the significant interactions between room condition and spatial separation [ $F(1,36)=91.2$ ,  $p < 0.001$ ], room condition and hearing loss [ $F(1,36)=9.0$ ,  $p = 0.005$ ],

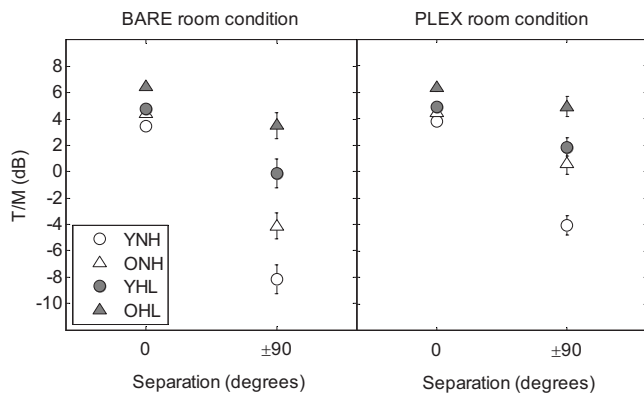


FIG. 2. Target-to-masker (T/M) ratio (expressed in decibels and computed relative to the level of individual maskers) at threshold in the two separation conditions ( $0^\circ$  and  $\pm 90^\circ$ ) for each listener group. The error bars are  $\pm 1$  standard error of the mean. The left panel contains results from the low-reverberant (BARE) room condition and the right panel shows results from the more reverberant (PLEX) room condition. Hearing status and age are denoted as shown in the symbol key.

and room condition, spatial separation, and hearing loss [ $F(1,36)=16.7$ ,  $p<0.001$ ]. Overall, the T/Ms at threshold were similar across room conditions when the talkers were colocated. Thresholds were at a target level that was slightly higher than that of the combined maskers (i.e., for a T/M at a threshold of 0 dB all talkers are equal in level whereas at +3 dB the target talker is 3 dB higher in level than each individual masker and approximately equal to the combined level of the two masker talkers). All eight mean thresholds (four groups by two rooms) occurred within a 2.9 dB range. For all groups, the average T/Ms at threshold when the talkers were spatially separated were lower than when they were colocated. The eight mean thresholds (four groups by two rooms) for the spatially separated talkers spanned an 11.7 dB range.

There were also significant interactions between spatial separation and hearing loss [ $F(1,36)=47.0$ ,  $p<0.001$ ] and spatial separation and age [ $F(1,36)=13.9$ ,  $p=0.001$ ], which are observable within each panel of Fig. 2. In each case, as pointed out above, the groups differed in performance very little for the colocated condition and much more for the spatially separated condition. All other interactions were nonsignificant. Despite the small range of performance in the colocated condition, within each room condition the groups were ordered YNH, ONH, YHL, and OHL going from the lowest to the highest T/M (average T/M at threshold in BARE: 3.5, 4.4, 4.8, and 6.4 dB and in PLEX: 3.8, 4.5, 4.9, and 6.3 dB, respectively). The same ordering of thresholds for the listener groups held for the much larger range of thresholds in the spatially separated condition (BARE: -8.2, -4.1, -0.1, and 3.5 dB; PLEX: -4.1, 0.6, 1.9, and 4.9 dB).

### C. Benefit of spatial separation

A comparison of the results for the colocated condition with those for the spatially separated condition was summarized in terms of the amount of spatial release from masking, a measure of the benefit of spatial separation. This value was calculated for each listener by subtracting the T/M at threshold for the  $\pm 90^\circ$  separation from the T/M at threshold for the

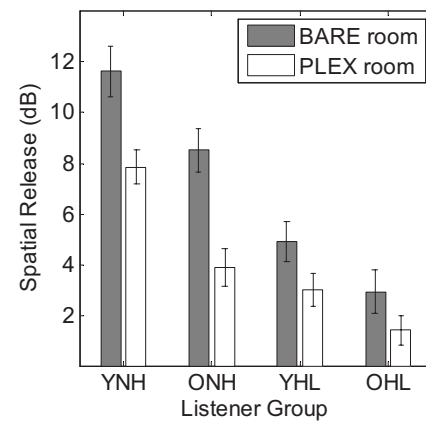


FIG. 3. Mean spatial release from masking in decibels (difference in T/M at threshold between colocated and separated conditions) for each listener group in both room conditions. The error bars are  $\pm 1$  standard error of the mean.

condition where the talkers were colocated. Figure 3 displays listener group on the abscissa with the amount of spatial release from masking in decibels along the ordinate. The mean and standard errors are shown for each listener group for both room conditions. This figure illustrates that all listener groups showed a benefit of spatial separation between the target and competing talkers in both room conditions. However, the amount of benefit was dependent on the listener group and on the room condition. Again, the order was consistent and held for both rooms with the largest release observed for the YNH group followed by ONH, YHL, and OHL groups. Statistical analysis by repeated-measures ANOVA on spatial release from masking was completed with room condition as a within-subjects factor and with hearing status and age as between-subjects factors. The results revealed significant main effects of room reverberation [ $F(1,36)=91.2$ ,  $p<0.001$ ], hearing status [ $F(1,36)=47.0$ ,  $p<0.001$ ], and age [ $F(1,36)=13.9$ ,  $p=0.001$ ]. There was a significant interaction between room reverberation and hearing status [ $F(1,36)=16.7$ ,  $p<0.001$ ] but not between room and age [ $F(1,36)=0.15$ ,  $p=0.70$ ], or room, hearing status, and age [ $F(1,36)=0.99$ ,  $p=0.33$ ].

Individual listener data are shown in Fig. 4 comparing the amount of spatial release in the BARE room condition with the amount of spatial release in the PLEX room condition. The solid line is a least-squares fit to the data and the dashed line shows equivalence of the two measures. Most data points fall below the dashed line consistent with the group results, indicating that the amount of spatial release in the BARE room was greater than in the PLEX room for most individuals. There was also a significant correlation between the amount of spatial release in the BARE room and the amount of spatial release in the PLEX room ( $r=0.86$ ,  $p<0.001$ ). Individual differences in spatial release can be observed within all listener groups. Those listeners with less spatial release overall showed less difference between the two room conditions (compare distances from the dashed line), which illustrates the source of the statistical interaction between hearing status and room reverberation also seen in the mean results. Note also the overlap between listener groups in the range of spatial release from masking: While



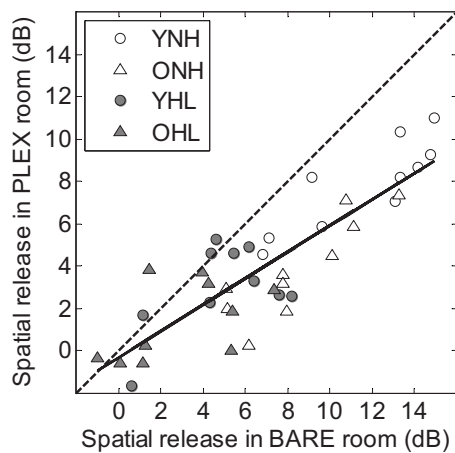


FIG. 4. Individual listener results for the amount of spatial release from masking in the PLEX room as a function of the amount of spatial release in the BARE room. The solid line is a least-squares fit to the data.

there was considerable overlap between younger and older listeners within each hearing status group there was much less overlap between listeners with and without hearing loss within each age group.

Figure 5 displays the relationship between spatial release from masking and speech identification thresholds in quiet (a measure related to the presence/magnitude of hearing loss). Results for the two room conditions are presented separately, with results for the BARE room condition in the left panel and results for the PLEX room condition in the right panel. The solid lines are least-squares fits to the data. Overall, the amount of spatial release from masking decreased with increasing quiet threshold. There was a strong and statistically significant negative correlation between quiet threshold and spatial release from masking in the BARE room condition ( $r = -0.84$ ,  $p < 0.001$ ). The slope was such that spatial release from masking decreased by about 2 dB for every 10 dB increase in speech identification threshold in quiet. This relationship was reduced to a moderate but significant correlation in the PLEX room condition ( $r = -0.66$ ,  $p < 0.001$ ). Because the listeners with the greatest amount of hearing loss were by necessity tested at a lower sensation level the influence of this factor should also be considered. Even if the six listeners who were tested at sensation levels below 20 dB in

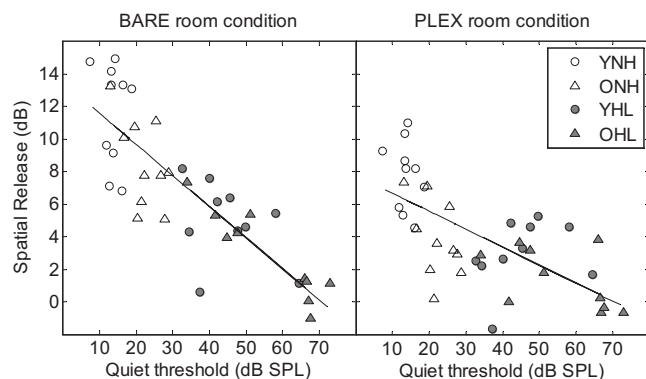


FIG. 5. Individual listener results showing the amount of spatial release from masking in decibels as a function of quiet threshold for identification of the CRM stimuli in each room condition. The solid lines are the least-squares fits to the data.

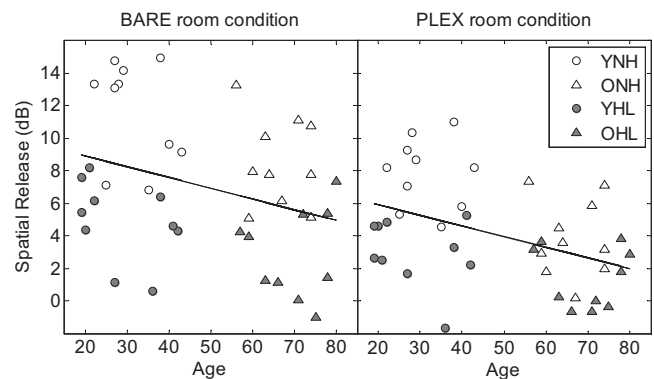


FIG. 6. Individual listener data for the amount of spatial release in decibels for each room condition as a function of age. The solid lines are the least-squares fits to the data.

the BARE room condition are excluded the correlation between quiet threshold and spatial release remains significant ( $r = -0.72$ ,  $p < 0.001$  in BARE). For the PLEX case, the correlation excluding the five listeners who were tested at sensation levels of 15 dB or below is  $-0.60$  ( $p < 0.001$ ). While it is prudent to keep in mind the potential confound of sensation level, these and other factors lead us to believe that it is not driving the effects seen here.<sup>2</sup>

The relationship between age and spatial release for individual listeners (coded by group) is shown in Fig. 6. The left panel displays the results in the BARE room condition and the right panel shows the results for the PLEX room condition. The solid lines are least-squares fits to the data. The negative correlations between age and spatial release were relatively weak but were significant in each room condition ( $r = -0.32$ ,  $p = 0.044$  in BARE;  $r = -0.43$ ,  $p = 0.005$  in PLEX). Because of the finding above, that just a 10 dB decrease in quiet threshold was associated with a 2 dB decrease in spatial release (for the BARE room condition), the difference in quiet thresholds across age groups needs to be taken into account. To help understand the relative contributions of these two variables, multiple regression analyses were performed with the amount of spatial release from masking for each room condition serving, in turn, as the predicted variable. Speech reception thresholds for the CRM stimuli in quiet and listener age were included as predictor variables. For the less reverberant room (BARE), 70% of the variance in the amount of spatial release from masking could be accounted for by variations in the speech identification threshold for the CRM stimuli in quiet ( $p < 0.001$ ). No significant increase in the variance accounted for by the model was achieved by the addition of age as a factor ( $p = 0.14$ ). For the more reverberant room (PLEX), the listener's speech reception threshold in quiet accounted for 44% of the variance ( $p < 0.001$ ) and including age as a factor increased the variance accounted for by 8% ( $p = 0.02$ ).

In summary, these results indicate that the listener's amount of hearing loss was the primary factor influencing the reduced benefit of spatial separation. This was less true as the amount of reverberation was increased. However, the effect of age, independent of hearing loss, appeared to show the opposite trend. The correlation and regression analyses



support a small, but statistically significant, effect of age on spatial release from masking in the more reverberant listening condition.

#### IV. DISCUSSION

The current study examined situations in which the listener must selectively attend to one talker in the presence of competing talkers. To do this, the listener must perceptually segregate the target speech stream and focus attention on it. There are several ways the listener could accomplish this task, including listening for the features of the talker's voice (e.g., fundamental frequency, prosody, and vocal tract length), using differences in level between talkers, following the content of the message, or using spatial cues. This study focused on the benefit of spatial information by holding all other factors constant or allowing them to vary in controlled but uninformative ways (e.g., randomizing talkers). In natural settings listeners typically monitor the entire auditory scene in order to switch attention to other sources when necessary, but for the current study they could focus exclusively on one spatial location. Because the target and masker sentences were spoken by same-sex talkers uttering highly similar and structured sentences, the task was subjectively quite difficult and made more so, we believe, by any degradation to the stimulus,<sup>3</sup> such as that caused by hearing loss and reverberation. The adverse effect of stimulus degradation was observed primarily in the spatially separated condition and that condition drove the magnitude of the difference underlying spatial release from masking. The similarity of the effect of hearing loss and reverberation on performance in the spatially separated condition suggests that other types of stimulus degradations (e.g., filtering, distorting, etc.) might likewise have an adverse effect on performance. However, the current results do not allow us to draw that conclusion and further study of the issue appears to be warranted.

##### A. Interpretation of colocated results

The threshold T/Ms for the target talker in the presence of two colocated masker talkers in the current study were comparable to the performance found by Brungart *et al.* (2001) for monaural segregation of three same-sex talkers using the same stimuli. The consistency of the thresholds across listeners and conditions (and studies, cf. Brungart *et al.*, 2001) was considered to be quite remarkable and suggests that the listeners have a reliable source segregation cue available. The most likely candidate, it seems to us, is the level of the target relative to the level of the maskers. According to this view, the presence of three simultaneous same-sex talkers obscures the talker-specific cues that the listener uses to follow the target voice from the call sign to the test words. The words that are reported are thus the more salient of those spoken. In this case, this occurs when the target level reaches a point where it is higher than the summed level of the two maskers (in our current metric the target level would equal the combined masker level at a T/M of +3 dB). Neither hearing loss nor reverberation appear to substantially affect this cue. This finding is in contrast to those of some other studies using only a single colocated

masker talker in which listeners with hearing loss typically required a higher T/M than listeners with normal hearing for equivalent performance (cf. 12.8 dB higher in Duquesnoy, 1983; approximately 5 dB higher in Gelfand *et al.*, 1988; and 3.7 dB higher in Arbogast *et al.*, 2005). However, these studies primarily tested older adults with hearing loss and in the current study, that group tended to require a slightly higher T/M than the other listener groups (see Fig. 2). In the study by Kidd *et al.* (2005) using methods similar to those used in the present study, threshold T/Ms in the PLEX room condition were 7 dB higher than in the BARE room condition. The main difference between studies was that there was only one masker talker and the stimuli were processed into nonoverlapping narrow frequency bands. The processing of the speech into narrow mutually exclusive frequency bands substantially reduced the spectral overlap of target and masker (reducing energetic masking, pitch, and intonation as factors) and may have provided a strong timbre cue useful in source segregation. Because of those differences, perhaps, the threshold T/Ms for both colocated and separated conditions were considerably lower than here. Thus we think it is likely that the difference between the findings of the current study and previous studies may be attributed to a difference in the segregation cues available for distinguishing between two versus three speech sources. The possible influence of a level "boundary" where a change in strategy may improve performance has been suggested in other studies of informational masking using a variety of different paradigms (e.g., Richards and Neff, 2004; Arbogast *et al.*, 2005).

##### B. Interpretation of spatially separated results

The results of the current experiment help to clarify one reason why listeners with hearing loss report difficulty selectively attending to speech in a multitalker background: They appear not to be able to make use of spatial separation cues as well as the listeners with normal hearing. In fact, for some listeners with hearing loss, their performance when the talkers were spatially separated was no better than when the talkers were colocated. The normal-hearing listeners were able to reach negative T/Ms at threshold; that is, they could tolerate the interfering talkers when they were actually substantially higher in level than the attended target talker. However, the listeners with hearing loss needed a higher, typically positive, T/M to understand the target talker correctly 50% of the time. The practical implication of this finding is important since listeners with hearing loss would likely require higher T/Ms than listeners with normal hearing to communicate effectively in everyday listening situations. As a result, the listeners with hearing loss had significantly less spatial release from masking than the normal-hearing listeners since the colocated condition did not differ much across listener groups. There was a strong and significant negative correlation between the amount of hearing loss and spatial release from masking. The relationship was such that the higher the listener's quiet threshold (i.e., more hearing loss), the less they benefited from the spatial separation cues. In these conditions, we speculate that factors that degrade the

stimulus, in this case hearing loss and reverberation, affect the performance more for spatially separated sources than for colocated sources.

The magnitude of the benefit of spatial separation between sources in the current study was larger for the listeners with normal hearing than in previous studies using symmetrically placed noise maskers and the subsequent difference between listener groups was also larger. [Bronkhorst and Plomp \(1992\)](#) used speech modulated noise maskers in a variety of symmetric or asymmetric configurations that were recorded in an anechoic room and presented to the listeners under headphones either monaurally or binaurally. In the conditions that are most relevant to the current experiment (two noise maskers either colocated with the target at 0° or symmetrically separated at  $\pm 90^\circ$  presented binaurally), spatially separating the maskers led to a 4.6 dB improvement in T/M for the normal-hearing listeners and a 2.7 dB improvement in T/M for the listeners with hearing loss. [Ter-Horst et al. \(1993\)](#) also found that listeners with hearing loss show less benefit than listeners with normal hearing from the separation of speech from two correlated noises shaped to match the long-term average speech spectrum. They tested symmetric configurations for both groups of listeners at  $\pm 18^\circ$  and  $\pm 54^\circ$  either horizontal or vertical separation. While the listeners with normal hearing had 4–6 dB spatial release for the horizontal symmetric separations, the listeners with hearing loss had around 1 dB spatial release on average.

Another study important for comparison to the current results is that of [Peissig and Kollmeier \(1997\)](#). Again, the trend in performance was similar to the current results but the size of the effect was smaller, which may be attributed to the emphasis on informational masking in the design of the current study. [Peissig and Kollmeier \(1997\)](#) compared the subjective intelligibility of two known target sentences in listeners with hearing loss to listeners with normal hearing in a virtual soundfield. In the condition using interfering speech tested in both listener groups, there were two fixed interfering sources (105° and 255°) and one variable interfering source (from 0° to 360°). For the five listeners with hearing loss that could be tested in this condition, they judged the target sentences to be 50% intelligible at signal-to-noise ratios that ranged from 0.6 to 3.8 dB lower when the variable interferer was spatially separated as compared to when it was colocated with the target. This benefit was less than that achieved by the normal-hearing listeners in the same condition who showed a maximum benefit of approximately 5 dB. Our results had the same pattern: Listeners with hearing loss showed a benefit of spatial separation between 0 and 8.2 dB while listeners with normal hearing obtained benefits up to a maximum of about 15 dB. There are, however, several important differences in the design of the two studies that could help to explain the differences in the size of the effects. First, the interfering sources in the current study were highly similar talkers uttering sentences that could plausibly be misidentified as targets. Second, there was also greater uncertainty in our procedure because the target talker and keywords varied from trial to trial, as did that of the other masker talkers. Finally, [Peissig and Kollmeier \(1997\)](#) used different-sex talk-

ers (two male and two female) providing the opportunity for segregating sources unrelated to spatial separation.

### C. Interpretation of the effect of hearing loss

Why should sensorineural hearing loss affect the ability of listeners to selectively attend to one talker in the presence of competing talkers? There are several possibilities that may not be mutually exclusive. A few of these are discussed here. It should be mentioned at the outset that achieving low thresholds for the task used in this study inherently depends on binaural information because no spatial release is obtained for monaural listening. As in the preceding study ([Marrone et al., 2008](#)), our interpretation of the spatial release obtained in these conditions is that it is primarily due to source segregation and focus of attention at a point in space that are enhanced by binaural cues.

First, it is possible that hearing loss increases the amount of energetic masking that the listeners experience. The CRM materials and testing procedure are designed to emphasize informational masking (e.g., [Brungart et al., 2001](#); [Arbogast et al., 2002](#)). Recently, [Brungart et al. \(2006\)](#) concluded that indeed informational masking normally dominates the masking observed in multitalker masking situations tested using the CRM. This conclusion was based on an analysis of the signal-to-noise ratio in spectrotemporal bins at various stimulus levels and the associated listener identification performance. However, the reduced frequency and temporal resolution associated with sensorineural hearing loss could produce greater energetic masking (spectrotemporal overlap) than observed in listeners with normal hearing. In support of that idea, [Arbogast et al. \(2005\)](#) found that there was greater energetic masking in listeners with hearing loss than in listeners with normal hearing. This difference was emphasized by the use of processing that rendered the target and masker speech into mutually exclusive sets of narrow bands. By comparing the masking produced by both narrow bands of noise and narrow bands of speech, they concluded that the reduced frequency resolution in the hearing-impaired listeners increased the ratio of energetic to informational masking. Because perceptual segregation cues are generally less effective for energetic masking than for informational masking (cf. [Freyman et al., 1999, 2001](#); [Kidd et al., 2008](#)), less spatial release from masking would be an expected consequence of increased energetic masking.

A second possible explanation for the decreased benefit of spatial separation of sources observed in listeners with hearing loss is a reduction in the extent to which the sound sources are perceptually segregated when they are spatially separated. According to this argument, for the listeners with hearing loss, the images of the stimuli may be less “distinct” than for the normal-hearing listeners. This assumes that segregation is not an all or none phenomenon but instead varies in strength depending on the potency of the cues used to form different sound images. A related issue has to do with the usefulness of ILDs in sound segregation. Although there is no usable better ear, and the current results are not consistent with the hypothesis discussed above regarding reduced spatial release because of reduced better-ear listening, there

is the possibility that the moment-to-moment differences in level among the sources—indicating different source locations—aid in segregation. The observation that listeners with hearing loss usually lose more high frequency information where ILDs are most potent could mean that that binaural cue for sound segregation is compromised. Although the monaural control condition tested by [Marrone et al. \(2008\)](#) rules out monaural “glimpsing” of the target during epochs of high T/M, it is not conclusive with respect to the possible role of time-varying ILDs in providing segregation cues. Thus, resolution of this issue awaits further study.

There is evidence from other studies using different paradigms that listeners with hearing loss may have poorer source segregation abilities as compared to those with normal hearing. Using alternating tone sequences to test sequential stream segregation, [Rose and Moore \(1997\)](#) found that one-half of their listeners with bilateral hearing loss judged perceptual fission of the two streams when the frequency difference between them was much larger than that needed by their normal-hearing counterparts. In a similar paradigm, [Mackersie et al. \(2001\)](#) related the boundary for perceptual fusion of two tone sequences to the perception of simultaneous sentences. They found that listeners who judged that perceptual fusion occurred at smaller frequency differences tended to have better understanding of two simultaneous sentences. [Grose and Hall \(1996\)](#) measured sequential processing on a gap detection/discrimination task and a melody recognition task. While the listeners with hearing loss showed the same patterns of performance on the tasks as those with normal hearing, they performed more poorly on both tasks overall. This was interpreted as evidence that hearing loss affects the perceptual organization of sequential stimuli, possibly due to a need for greater frequency separation between auditory streams as compared to listeners with normal hearing. In a different paradigm, [Kidd et al. \(2002\)](#) found poorer perceptual segregation of individual components of complex sounds as a function of increasing amount of hearing loss. They measured detection of a pure tone sequence embedded in a multiple burst masker (either energetic or informational) as well as auditory filter measurements. The listeners with hearing loss showed a reduced ability to detect the target in a spectrotemporally complex and uncertain masker, a result which could not be explained by their auditory filter characteristics. This finding was instead attributed to a difference in analytic listening abilities, a conclusion that warrants consideration in the current study where analytic listening abilities are critical for hearing out one voice among several similar competing voices.

Finally, differences in other higher-level processes important to performing the task—such as attentional focus—between the groups could drive the observed differences in the benefit of spatial separation. If we assume that the different categories of listeners have the same pool of cognitive/attentional resources available, hearing loss may place an additional processing load on the observer leaving a smaller residual pool of resources available to solve the task. Consequently, if this conjecture were true, listeners with hearing loss would be predicted to perform equivalently to normal-hearing listeners when the selective listening task was rela-

tively easy but would perform poorly when the task was difficult and required more processing resources. It seems reasonable to speculate that the listening environment becomes more complex and demanding when the number of distinct sources, the degree of listener uncertainty, and the distraction value of the sources increase. In the present study, all of these factors were, by design, chosen to be relatively high. A related but somewhat different explanation is that the listeners with hearing loss simply have difficulty inhibiting the processing of irrelevant stimuli. For example, [Doherty and Lutfi \(1999\)](#) found that listeners with sensorineural hearing loss had greater difficulty ignoring unwanted sounds as compared to normal-hearing listeners in a level discrimination task. Presumably, the greater the similarity of the target and maskers, the greater the demands on cognitive processing and the more difficult it is to ignore irrelevant stimuli. In the current experiment, the listeners with hearing loss informally reported that the task was quite effortful and often stated that this was because they had difficulty ignoring the masker talkers even when they were spatially separated.

#### D. Effect of reverberation

For all listener groups, performance in the spatially separated condition was also adversely influenced by reverberation. When both hearing loss and reverberation were present, performance was worse than with either alone. There was a relationship between the benefit of spatial separation observed in the more reverberant room and how much benefit was shown in the less reverberant room. Listeners demonstrating larger spatial release in the BARE room condition were those listeners with the larger spatial release in the PLEX room condition. These listeners also showed the greatest difference between the room conditions (i.e., greatest deviation from the diagonal line in Fig. 4). The present results differ somewhat from similar measures in normal-hearing listeners reported by [Kidd et al. \(2005\)](#). In that study, increased reverberation elevated speech identification thresholds in *both* colocated and spatially separated conditions about equally, so the large spatial release they found was preserved. As noted above, we believe that the reduction in spatial release found here is a consequence of the limited utility of normal talker segregation cues in the colocated condition leaving relative level as the primary, and robust, basis for selecting the target talker.

#### E. Effect of age

Although age was a significant factor in the overall results from the repeated-measures ANOVAs, this must be interpreted with caution due to the results of further analyses. The small difference in thresholds between age groups (cf. Fig. 1) is enough to account for part of the difference in spatial release as indicated by the lines representing the least-squares fits shown in Fig. 5. Furthermore, the correlations between spatial release and age were weak (cf. Fig. 6) and the regression analyses indicate no further contribution of age once quiet threshold is accounted for in the less reverberant room and only a small contribution (an additional 8% of the variance accounted for) in the more reverberant room.



Consequently, performance differences observed across listener groups cannot unambiguously be attributed to age effects independent of hearing status. This is consistent with the finding of Li *et al.* (2004), where younger and older adults with clinically normal hearing were tested in conditions in which the target and masker talkers were perceived to be originating from different locations (cf. Freyman *et al.*, 1999). Overall, the relative improvement due to perceived spatial separation was the same for younger and older adults, although the older adults needed a slightly higher T/M for equivalent performance. Dubno *et al.* (2008) also found that the older adults had generally higher thresholds but the same amount of benefit due to spatial separation for the speech in noise task using symmetric maskers. However, that benefit was smaller than their articulation-index based predictions. Interestingly, Humes *et al.* (2006) found that older adults performed significantly worse than younger adults on eight of nine divided attention tasks but on only three of nine selective attention tasks. This suggests that an age effect might be observable under the conditions of the current experiment if the task demands were more challenging; for example, if the keywords from two talkers must be reported on every trial. Recently, Helfer and Freyman (2008) compared speech-on speech masking in younger and older listeners under conditions in which differences in apparent spatial location of target and masker normally provide a significant enhancement in performance. Although several types of maskers were tested, of particular interest here was their finding that older listeners were less able to use spatial cues to improve speech recognition in general but especially for different-sex, rather than same-sex, talkers. This effect was particularly apparent in older listeners with hearing loss. Thus, a greater age effect might have occurred in this study had different-sex talkers been tested. The authors conclude that age-related changes in cognitive function may affect the ability to selectively attend to one talker among other spatially separated talkers, an effect that may be related to a general difficulty in ignoring irrelevant stimuli of high semantic content.

## V. SUMMARY

Consistent with the report by Marrone *et al.* (2008), normal-hearing listeners demonstrated a large benefit of spatial separation when listening to multiple simultaneous talkers and trying to identify the speech of one of them. This effect was obtained without the availability of a better-ear advantage and was relatively robust with respect to increased reverberation. Bilateral sensorineural hearing loss decreased the benefit of spatial separation of sources. All 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 much higher T/Ms at threshold than normal-hearing listeners, particularly in the reverberant environment. For some listeners with hearing loss, performance was as poor in the spatially separated condition as in the colocated condition. The effect of age separate from hearing loss, however, was inconclusive. Increased energetic masking and decreased ability to segregate

and focus attention in space were considered possible explanations for the findings. Overall, these results help to explain why listeners with hearing loss often report great difficulty understanding a target talker in real rooms when there are other talkers in the background.

## ACKNOWLEDGMENTS

This work was supported by Grant Nos. DC004545, DC00100, and DC004663 from the National Institute on Deafness and Other Communication Disorders (NIDCD). N.M. was supported by a Ruth L. Kirschstein National Research Service Award for predoctoral fellows from NIDCD (F31 DC008440). This work partly fulfills the requirements for the Doctor of Philosophy degree at Boston University by N.M. We are grateful to Deborah Corliss and Janine Francis for their assistance in the Sound Field Laboratory. Thanks also to Virginia Best, Barbara Shinn-Cunningham, Melanie Matthies, Steven Colburn, Richard Freyman, and three anonymous reviewers for helpful comments on earlier versions of this manuscript.

<sup>1</sup>The measurements as described in Kidd *et al.*, 2005 pertain directly here as they involve the exact same surface treatments, loudspeakers and loudspeaker positions, listener position, etc., as in the current experiment. The purpose was to compare across our own experimental room conditions; therefore, generalization to other rooms and different measurement procedures (e.g., omnidirectional source) should be made with caution.

<sup>2</sup>To obtain more insight into the possible effects of sensation level on spatial release from masking several listeners were retested at different sensation levels in the BARE room condition using methods identical to the main experiment. When four listeners from the YNH group were retested at 20 dB SL (equal to or less than that used with most listeners in the hearing loss groups), they continued to show large spatial release from masking. In addition, when two listeners from the YHL group were tested at higher sensation levels their spatial release from masking did not improve. However, when the four YNH listeners and four additional YHL listeners were tested at 10 dB SL, the amount of spatial release was reduced but not eliminated on an individual listener basis within both groups. Although these individuals may not be representative of the larger groups we believe that this in addition to the fact that the correlations remain significant between quiet threshold and spatial release when the low SL listeners are excluded is convincing evidence that the different SLs used are not the reason for the effect.

<sup>3</sup>One effect of reverberation is the disruption of the binaural cues of ILD and ITD, which we include in our characterization of stimulus degradation. As discussed in the text, cochlear hearing loss also may be thought of as causing a degradation of the stimulus.

American National Standards Institute. (2004). American national standard specification for audiometers. ANSI 3.6-2004. New York, NY: Acoustical Society of America.

Arbogast, T. L., Mason, C. R., and Kidd, G., Jr. (2002). "The effect of spatial separation on informational and energetic masking of speech," *J. Acoust. Soc. Am.* **112**, 2086–2098.

Arbogast, T. L., Mason, C. R., and 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**, 2169–2180.

Best, V., Gallun, F. J., Ihlefeld, A., and Shinn-Cunningham, B. G. (2006). "The influence of spatial separation on divided listening," *J. Acoust. Soc. Am.* **120**, 1506–1516.

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. W., and Plomp, R. (1989). "Binaural speech intelligibility in noise for hearing-impaired listeners," *J. Acoust. Soc. Am.* **86**, 1374–1383.

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

- Bronkhorst, A. (2000). "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acust. Acta Acust.* **86**, 117–128.
- Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). "Informational and energetic masking effects in the perception of multiple simultaneous talkers," *J. Acoust. Soc. Am.* **110**, 2527–2538.
- Brungart, D. S., and Simpson, B. D. (2002). "Within-ear and across-ear interference in a cocktail-party listening task," *J. Acoust. Soc. Am.* **112**, 2985–2995.
- Brungart, D. S., Chang, P. S., Simpson, B. D., and Wang, D. (2006). "Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation," *J. Acoust. Soc. Am.* **120**, 4007–4018.
- Colburn, H. S., Shinn-Cunningham, B. A., Kidd, G., Jr., and Durlach, N. I. (2006). "The perceptual consequences of binaural hearing," *Int. J. Audiol.* **45**, 34–44.
- Doherty, K. A., and Lutfi, R. A. (1999). "Level discrimination of single tones in a multitone complex by normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **105**, 1831–1840.
- Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (2002). "Spectral contributions to the benefit from spatial separation of speech and noise," *J. Speech Lang. Hear. Res.* **45**, 1297–1310.
- Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (2008). "Binaural advantage for younger and older adults with normal hearing," *J. Speech Lang. Hear. Res.* **51**, 539–556.
- Duquesnoy, A. J. (1983). "Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons," *J. Acoust. Soc. Am.* **74**, 739–743.
- Ebata, M. (2003). "Spatial unmasking and attention related to the cocktail party problem," *Acoust. Sci. & Tech.* **24**, 208–219.
- Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). "The role of perceived spatial separation in the unmasking of speech," *J. Acoust. Soc. Am.* **106**, 3578–3588.
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2001). "Spatial release from informational masking in speech recognition," *J. Acoust. Soc. Am.* **109**, 2112–2122.
- Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2004). "Effect of number of masking talkers and priming on informational masking in speech recognition," *J. Acoust. Soc. Am.* **115**, 2246–2256.
- Gatehouse, S., and Noble, W. (2004). "The speech, spatial, and qualities of hearing scale (SSQ)," *Int. J. Audiol.* **43**, 85–99.
- Gelfand, S. A., Ross, L., and 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.
- Grose, J. H., and Hall, J. W. (1996). "Perceptual organization of sequential stimuli in listeners with cochlear hearing loss," *J. Speech Hear. Res.* **39**, 1149–1158.
- Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (1999). "Speech intelligibility and localization in complex environments," *J. Acoust. Soc. Am.* **105**, 3436–3448.
- Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," *J. Acoust. Soc. Am.* **115**, 833–843.
- Helfer, K. S. (1992). "Aging and the binaural advantage in reverberation and noise," *J. Speech Hear. Res.* **35**, 1394–1401.
- Helfer, K. S., and Freyman, R. L. (2008). "Aging and speech-on-speech masking," *Ear Hear.* **29**, 87–98.
- Humes, L. E., Lee, J. H., and Coughlin, M. P. (2006). "Auditory measures of selective and divided attention in young and older adults using single-talker competition," *J. Acoust. Soc. Am.* **120**, 2926–2937.
- Kidd, G., Jr., Mason, C. R., Rohdla, T. L., and Deliwalla, P. S. (1998). "Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns," *J. Acoust. Soc. Am.* **104**, 422–431.
- Kidd, G., Jr., Arbogast, T. L., Mason, C. R., and Walsh, M. (2002). "Informational masking in listeners with sensorineural hearing loss," *J. Assoc. Res. Otolaryngol.* **3**, 107–119.
- Kidd, G., Jr., 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. (2008). "Informational masking" in *Auditory Perception of Sound Sources*, edited by W. A. Yost (Springer, New York), pp. 143–189.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Li, L., Daneman, M., Qi, J. G., and Schneider, B. A. (2004). "Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults?," *J. Exp. Psychol. Hum. Percept. Perform.* **30**, 1077–1091.
- Mackersie, C. L., Prida, T. L., and Stiles, D. (2001). "The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss," *J. Speech Lang. Hear. Res.* **44**, 19–28.
- Marrone, N., Mason, C. R., and Kidd, G., Jr. (2008). "Tuning in the spatial dimension: Evidence from a masked speech identification task," *J. Acoust. Soc. Am.* **124**, 1146–1158.
- Noble, W., and Perrett, S. (2002). "Hearing speech against spatially separate competing speech versus competing noise," *Percept. Psychophys.* **64**, 1325–1326.
- Peissig, J., and Kollmeier, B. (1997). "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," *J. Acoust. Soc. Am.* **101**, 1660–1670.
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). "How young and old adults listen to and remember speech in noise," *J. Acoust. Soc. Am.* **97**, 593–608.
- Richards, V. M., and Neff, D. L. (2004). "Cuing effects for informational masking," *J. Acoust. Soc. Am.* **115**, 289–300.
- Rose, M. M., and Moore, B. C. (1997). "Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **102**, 1768–1778.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Shinn-Cunningham, B. G., Ihlefeld, A., Satyavarta, and Larson, E. (2005). "Bottom-up and top-down influences on spatial unmasking," *Acust. Acta Acust.* **91**, 967–979.
- Ter-Horst, K., Byrne, D., and Noble, W. (1993). "Ability of hearing-impaired listeners to benefit from separation of speech in noise," *Aust. J. Audiol.* **15**, 71–84.
- 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 (Erlbaum, Mahwah, NJ), pp. 329–348.
- Zurek, P. M. (1993). "Binaural advantages and directional effects in speech intelligibility" in *Acoustical Factors Affecting Hearing Aid Performance*, 2nd ed., edited by G. A. Studebaker and I. Hockberg (Allyn and Bacon, Needham Heights, MA), pp. 255–276.