

# Effect of audibility on spatial release from speech-on-speech masking<sup>a)</sup>

Helen Glyde<sup>b),c)</sup> and Jörg M. Buchholz<sup>b),d),e)</sup>

The HEARing Cooperative Research Centre, 550 Swanston Street, Carlton, Victoria 3010, Australia

Lillian Nielsen, Virginia Best,<sup>f)</sup> Harvey Dillon,<sup>g)</sup> and Sharon Cameron<sup>g)</sup>

National Acoustic Laboratories, 16 University Avenue, Macquarie University, Sydney, New South Wales 2067, Australia

Louise Hickson<sup>g)</sup>

School of Health and Rehabilitation Sciences, The University of Queensland, St Lucia, Queensland 4072, Australia

(Received 25 August 2014; revised 1 October 2015; accepted 15 October 2015; published online 25 November 2015)

This study investigated to what extent spatial release from masking (SRM) deficits in hearing-impaired adults may be related to reduced audibility of the test stimuli. Sixteen adults with sensorineural hearing loss and 28 adults with normal hearing were assessed on the Listening in Spatialized Noise–Sentences test, which measures SRM using a symmetric speech-on-speech masking task. Stimuli for the hearing-impaired listeners were delivered using three amplification levels (National Acoustic Laboratories - Revised Profound prescription (NAL-RP) +25%, and NAL-RP +50%), while stimuli for the normal-hearing group were filtered to achieve matched audibility. SRM increased as audibility increased for all participants. Thus, it is concluded that reduced audibility of stimuli may be a significant factor in hearing-impaired adults' reduced SRM even when hearing loss is compensated for with linear gain. However, the SRM achieved by the normal hearers with simulated audibility loss was still significantly greater than that achieved by hearing-impaired listeners, suggesting other factors besides audibility may still play a role.

© 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4934732>]

[FJG]

Pages: 3311–3319

## I. INTRODUCTION

Spatial cues are some of the most salient cues available to listeners to aid in the segregation of speech streams in noisy environments. Typically, in real world situations, target speech and interfering noise arise from different directions. The resulting differences in interaural time differences (ITDs) and interaural level differences (ILDs) make separating the speech from the noise easier (Bamiou, 2007). The ability to exploit these spatial cues has been referred to in the literature by many terms including spatial release from masking (SRM), spatial unmasking, and spatial processing, and is commonly measured as the difference in intelligibility between a condition where target and noise sources are

spatially separated and a reference condition where all sources are co-located. In many cases, a spatial advantage arises because the physical locations of the sounds offer an improvement in the effective signal-to-noise ratio (SNR). An improved SNR can reduce both “energetic masking” (EM), where sounds compete for representation at the auditory periphery, and “informational masking” (IM), where sounds are confusable and compete for central resources (see Kidd *et al.*, 2008 for a review). When IM is the dominant problem, it appears that the *perceived* separation of sources, even in the absence of an improved SNR, can provide an advantage by enabling attention to be directed selectively (e.g., Freyman *et al.*, 1999).

Normal-hearing (NH) adults are able to gain as much as 20 dB of benefit from SRM (Bronkhorst and Plomp, 1988), although the amount of benefit varies depending on the physical properties of the interfering noise, the relative amounts of EM and IM, the degree of spatial separation, and the listener's task. Unfortunately, the same degree of benefit has not been seen in hearing-impaired (HI) listeners (e.g., Gelfand *et al.*, 1988; Dubno *et al.*, 2002; Arbogast *et al.*, 2002; Marrone *et al.*, 2008; Glyde *et al.*, 2013c). There are several broad hypotheses that have been considered as to why HI listeners show reduced SRM. First, it is possible that reduced audibility limits the availability of speech information, as well as the associated spatial cues, both of which reduce the potential for SRM. This is arguably the simplest hypothesis and is of primary interest in this study.

<sup>a)</sup>Sections of the research reported here were also reported in the doctoral dissertation of H.G. and the masters dissertation of L.N. Findings from this research were presented at the World Congress of Audiology XXXII (Brisbane, Australia, 2014) and at the Midwinter meeting of the ARO (San Diego, USA, 2014).

<sup>b)</sup>Also at the National Acoustic Laboratories, 16 University Avenue, Macquarie University, Sydney, NSW 2067, Australia.

<sup>c)</sup>Also at the School of Health and Rehabilitation Sciences, The University of Queensland, St Lucia, QLD 4072, Australia.

<sup>d)</sup>Also at the Department of Linguistics, 16 University Avenue, Macquarie University, Sydney, NSW 2067, Australia.

<sup>e)</sup>Electronic mail: [jorg.buchholz@nal.gov.au](mailto:jorg.buchholz@nal.gov.au)

<sup>f)</sup>Also at the Department of Speech, Language and Hearing Sciences, Boston University, Boston, MA 02215, USA.

<sup>g)</sup>Also at the HEARing Cooperative Research Centre, 550 Swanston Street, Carlton, VIC 3010, Australia.

Alternative hypotheses that have been raised include reductions in the fidelity of spatial cues as a result of degraded neural coding, a change in the relative amounts of EM and IM, or a reduced ability to direct spatial attention.

We consider three distinct ways in which reduced audibility has the potential to affect SRM. First, there will be a reduced dynamic range for absolute target audibility, which might limit the amount of improvement in intelligibility that is possible given spatial separation. Second, loss of audibility could prevent HI people from accessing improvements in SNR in the “better ear” usually provided by the head shadow when competing sounds are spatially separated. As the head shadow is frequency-dependent, these improvements in SNR are primarily available at high frequencies (i.e., above about 1.5 kHz) where typical hearing losses are most severe. Third, whatever components of the target and masker that are perceptible will be received at a reduced sensation level, potentially reducing their salience and that of the ITD and ILD cues associated with them.

Despite decades of research investigating spatial processing ability in HI people the role played by reduced audibility of the test stimuli, particularly for paradigms that use speech maskers, remains uncertain. One test configuration that has been used in numerous recent studies involves the presentation of speech maskers separated symmetrically to either side of the head (Marrone *et al.*, 2008; Best *et al.*, 2012; Gallun *et al.*, 2013; Glyde *et al.*, 2013b; Glyde *et al.*, 2013c). This configuration produces a large SRM in NH listeners, but a much smaller SRM in HI listeners. While all listeners tend to perform similarly (and poorly) in the collocated configuration, NH listeners are able to achieve much lower thresholds in the symmetric masker configuration. For this paradigm, the robust SRM in NH listeners has been attributed largely to the perceived separation causing a release from IM, although it has also been shown that the size of the release can be related to energetic factors or the opportunity for better-ear glimpsing (Brungart and Iyer, 2012; Glyde *et al.*, 2013a).

Previous studies have employed flat amplification (at the source) for HI participants in an attempt to remove audibility as a confounding factor (e.g., Marrone *et al.*, 2008). Though this approach improves overall audibility, it does not account for the sloping nature of most hearing losses and thus does not provide “normal” audibility across the spectrum. In our previous experiment using the Listening in Spatialized Noise–Sentences test (LiSN-S; Glyde *et al.*, 2013c) individualized linear amplification according to the revised National Acoustic Laboratories prescription procedure for profound hearing losses (NAL-RP) was utilized to compensate for reduced audibility. Despite this, SRM was negatively correlated with degree of hearing loss (partial  $r^2 = 0.66$ ) when age was controlled for. For instance, a person with a 60 dB four-frequency average hearing loss (4FAHL) gained only 4 dB of benefit from spatial separation as compared with 14 dB of benefit seen for NH listeners. NAL-RP, like most fitting prescriptions, does not attempt to provide audibility equal to that experienced by NH listeners (i.e., it does not “invert the audiogram”) (Dillon, 2012). By providing amplification we can be confident that the target

speech was intelligible to the HI listeners in quiet, however we cannot assert they had the same amount of access to the high-frequency information as NH listeners. Furthermore, NAL-RP was developed to provide optimal gain for a 65 dB sound pressure level (SPL) input but the levels of the LiSN-S stimuli are lower than this. The fixed-level maskers are set at a combined level of 55 dB SPL, which would have resulted in an output that was lower than optimal. Therefore, it is possible that reduced audibility was a contributing factor to reduced SRM seen in HI listeners in Glyde *et al.* (2013c). One way to evaluate the effect of audibility is to provide extra amplification above that provided by NAL-RP. However, the amount of extra amplification that can be applied to an individual subject is limited by loudness discomfort, which typically does not allow full restoration of NH audibility in HI subjects. An alternative, or complementary, method is to simulate hearing loss by reducing audibility in NH subjects. This experimental approach has the advantage that reduced audibility can be studied in isolation, since NH listeners do not have other deficits, commonly associated with hearing loss, that may also affect SRM.

The current study combined these two approaches to evaluate the effect of audibility on SRM. To the extent possible, different levels of extra amplification were applied on top of the NAL-RP prescription to HI adults. Additionally we simulated the audibility resulting from the combination of hearing loss and these different amplification levels in NH listeners by filtering (i.e., attenuating) the speech material. It was hypothesized that HI listeners provided with additional gain on top of NAL-RP would show an increase in SRM as measured by the LiSN-S, and that the measured SRM would be similar in NH listeners tested with the corresponding filtered stimuli.

## II. EXPERIMENTAL MATERIAL AND METHODS

### A. Participants

The study was conducted under the ethical clearance and guidance of the Australian Hearing Ethics Committee. All participants provided written informed consent prior to testing and received a gratuity of \$20 to cover travel costs associated with attending the research appointment. A summary of demographic information for each of the participant groups and comparison samples are available in Table I.

#### 1. Hearing-impaired participants

Sixteen adults with a sloping mild-to-moderate sensorineural hearing loss, with an air-bone gap smaller than 10 dB, and who were aged between 21 and 80 years (mean = 68.8 years) participated in the study. Hearing losses were symmetrical, defined as left- and right-ear thresholds at octave frequencies from 250 to 4000 Hz being within 10 dB of each other. Larger differences were accepted at frequencies above 4 kHz and occurred in a number of tested subjects. Mean audiometric thresholds are shown in Table II. All participants were experienced hearing aid users (>2 years) and were required to have English as their first language to be eligible to participate. Moreover, they were all healthy, fully functional, independent volunteers with no reported

TABLE I. Summary table of participant details. The different participant groups and the applied amplification processing are further described in Secs. II A and II C, respectively.

	Number of Participants	Amplification	Mean Age $\pm$ 1 STD	Age Range
HI group	16	NAL-RP + 25% and NAL-RP + 50%	68.8 $\pm$ 14.3 years	21–80 years
Simulated HI group A	12	NAL-RP	33.6 $\pm$ 6.1 years	25–47 years
Simulated HI group B	16	NAL-RP + 25% and NAL-RP + 50%	28.8 $\pm$ 10.8 years	18–53 years
Comparison HI group Glyde <i>et al.</i> (2013c)	16	NAL-RP	73.1 $\pm$ 14.5 years	39–87 years
Comparison NH group Cameron <i>et al.</i> (2011)	96	N/A	31.9 $\pm$ 11.8 years	18–60 years

cognitive deficits and were taken from a database available at the National Acoustic Laboratories.

## 2. Normal-hearing participants

Twenty-eight NH adults aged between 18 and 53 years (mean = 30.9 years) also took part and made up the simulated HI groups. Each participant had hearing thresholds equal to or better than 20 dB hearing level (HL) at each octave frequency from 250 Hz to 8000 Hz which did not differ more than 10 dB between ears. The NH participants were divided into two groups for which different levels of reduced audibility were applied. Group A contained 12 NH adults who were given stimuli with an audibility that simulated hearing loss with NAL-RP amplification. The remaining 16 NH adults were assigned to group B and were given stimuli with an audibility simulating amplification levels of NAL-RP + 25% and NAL-RP + 50% on top of hearing loss.

## B. Speech material

As in Glyde *et al.* (2013c) participants were assessed on the LiSN-S. The LiSN-S was selected for use as it is a clinically standardized measure of SRM. The LiSN-S, which is described in detail in Cameron and Dillon (2007), assesses SRM using a sentence repetition task conducted under headphones. Stimuli are convolved with head related transfer functions (HRTFs) to provide simulated spatial cues. The test includes 120 short target sentences (e.g., “The brother carried her bag”) which are presented from 0° azimuth in the presence of two competing children’s stories which act as maskers. In the current study all participants were assessed on: the same voices 0° (SV0) condition in which the maskers are voiced by the same female speaker as the target sentences and are perceived as coming from 0° azimuth; and the same voices  $\pm 90^\circ$  (SV90) condition in which the maskers are voiced by the same female speaker as the target sentences, but one masker emanates from +90° azimuth while the other emanates from -90° azimuth. The amount of SRM

obtained by each participant is calculated as the difference in performance between the SV90 and SV0 conditions. This derived score is termed the spatial advantage. In order to allow more levels of audibility to be assessed without repeating sentence lists, the decision was made to only include the same voice conditions (which causes the most IM and also produces the largest difference between NH and HI listeners; Glyde *et al.*, 2013c). Sentence lists were counterbalanced across conditions.

## C. Audibility processing

In addition to that provided by NAL-RP, the present study measured two additional audibility levels. Since NAL-RP is only defined for a frequency range of up to 6 kHz, it was extended here by simply applying the standard NAL-RP gain prescription formula above 6 kHz (Dillon, 2012) while setting the required parameter  $k$  to -2 dB. Above 8 kHz the gain was kept constant and was lowpass filtered at 12 kHz by applying a 16th-order Butterworth filter. As further described in Secs. II C 1 and II C 2 an additional 25% and 50% of gain on top of NAL-RP prescription levels were used. NAL-RP + 50% was judged to be the maximum gain that could be presented safely to all participating HI listeners without encountering issues of discomfort. Details regarding how these audibility levels were achieved for both groups are provided in the following two sections.

### 1. Processing applied for hearing-impaired participants

The individual, frequency-dependent gains (in dB) that were applied to the speech mixture to provide different levels of audibility were calculated as

$$g_{HI} = g_{RP} + \alpha(HL - g_{RP}^*) \quad (1)$$

with HL the hearing loss of the considered test subject averaged across ears,  $g_{RP}$  the corresponding gain prescribed by NAL-RP (Sec. II C),  $\alpha$  a gain-factor  $0 \leq \alpha \leq 1$  that controls

TABLE II. The average hearing thresholds with  $\pm 1$  standard deviation for the hearing-impaired group, the comparison hearing-impaired sample from Glyde *et al.* (2013c) and the hearing loss simulated for the NH participants.

	Frequency (Hz)					
	250	500	1000	2000	4000	8000
HI participants’ thresholds (dB HL)	26.6 $\pm$ 3.9	35.5 $\pm$ 2.9	42.2 $\pm$ 2.8	50.8 $\pm$ 5.0	61.3 $\pm$ 4.2	73.0 $\pm$ 3.9
Comparison Sample’s thresholds (Glyde <i>et al.</i> 2013c) (dB HL)	27.0 $\pm$ 5.3	33.0 $\pm$ 4.5	43.0 $\pm$ 4.6	55.0 $\pm$ 3.4	63.0 $\pm$ 6.8	75.0 $\pm$ 2.6
Thresholds simulated for NH participants (dB HL)	28	32	42	55	62	75

the extra gain that is applied on top of the gain prescribed by NAL-RP, and  $g_{RP}^*$  an effective gain given by

$$g_{RP}^* = L_{HI} - L_{NH} + g_{RP} \quad (2)$$

with  $L_{NH}$  the speech level at the output of a NH auditory bandpass (BP) filterbank averaged over all target sentences and  $L_{HI}$  the corresponding speech level at the output of a HI auditory BP filterbank. In the case of  $\alpha=0$ , Eq. (1) prescribes a gain according to NAL-RP (i.e.,  $g_{HI} = g_{RP}$ ). In the case of  $\alpha=1$ , a very high gain is prescribed that fully restores NH audibility. The value of  $\alpha$  therefore indicates the extent to which the additional gain removes the deficit in audibility that occurs with NAL-RP amplification relative to the audibility provided by normal hearing. The frequency-dependent gains given in Eq. (1) were realized by linear-phase finite impulse response (FIR) filters with a length of 1024 samples at a sampling frequency of  $f_s = 44.100$  Hz.

The BP-filters applied in the derivation of the speech levels  $L_{HI}$  and  $L_{NH}$  were realized by fourth-order Gammatone filters (Patterson *et al.*, 1988). The bandwidth of the NH BP-filters was equal to one equivalent rectangular bandwidth (ERB: Patterson *et al.*, 1988). The HI BP-filters had an increased bandwidth, which was calculated according to the equations provided by Nejime and Moore (1997) and by applying the average hearing thresholds provided in Table II. The different filterbanks take into account the effect of spectral integration of the target signal by the auditory filters on audibility that is only relevant for broadband signals like speech but not for pure tones as used in an audiogram. With reference to Eq. (2), the broader auditory filters in the impaired system result in an increased effective speech level (i.e.,  $L_{HI} \geq L_{NH}$ ) and thus, result in an increased effective gain (i.e.,  $g_{RP}^* \geq g_{RP}$ ). For the above HI auditory filters this frequency-dependent increase is up to 6 dB.

Three different audibility levels were realized:  $\alpha=0$  (NAL-RP),  $\alpha=0.25$  (NAL-RP + 25%), and  $\alpha=0.50$  (NAL-RP + 50%). The resulting average (broadband) speech levels for the target sentences at an SNR of 0 dB (corresponding to a target level of 55 dB) were 74, 82, and 91 dB SPL. Figure 1 (left panel) shows the gains prescribed for the HL given in Table II for  $\alpha=0$  (NAL-RP),  $\alpha=0.50$

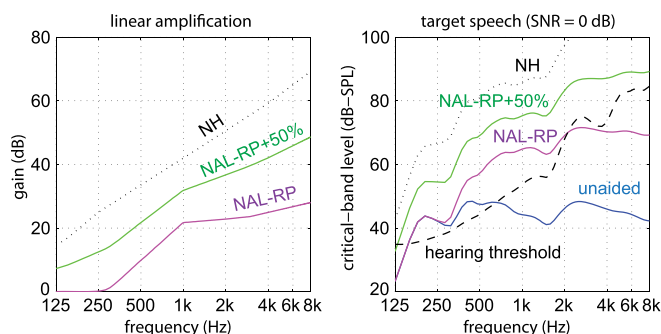


FIG. 1. (Color online) Example gains that were applied to realize different levels of audibility (left panel). The corresponding average target speech spectra at an SNR of 0 dB are shown in the right panel together with the considered hearing threshold (see Table II).

(NAL-RP + 50%), and  $\alpha=1$  (NH). The corresponding (amplified) speech spectra at the output of the described HI BP filterbank are shown in the right panel of Fig. 1 together with the relevant hearing threshold. From this figure it can be deduced that the different amplification levels have a direct effect on the audible bandwidth of the sentence materials.

## 2. Processing applied for normal-hearing participants

The speech materials were adjusted (attenuated) to provide audibility to a NH participant equal to that which would have been experienced by an aided HI participant with hearing thresholds equal to those shown in Table II. Based on the concepts described in Sec. II C 1, the required attenuation was realized by linear-phase FIR filters with a length of 1024 samples ( $f_s = 44.100$  Hz) and a gain response (in dB) given by

$$g_{Hsim} = (1 - \alpha)(g_{RP}^* - HL) \quad (3)$$

with HL the considered hearing loss,  $\alpha$  the gain-factor introduced in Sec. II C 1, and  $g_{RP}^*$  given in Eq. (2). In addition to the audibility levels described in Sec. II C 1 (i.e.,  $\alpha=0$ ,  $\alpha=0.25$ , and  $\alpha=0.50$ ) the standard NH condition was also realized (i.e.,  $\alpha=1$ ). This resulted in average (broadband) speech levels for the target sentences at an SNR of 0 dB of 31, 37, 43, and 55 dB SPL.

To illustrate the processing involved in the described hearing loss simulations, the target speech spectra for the average HI subject at the output of a HI BP filterbank are replotted in Fig. 2(a) from Fig. 1 (left panel) for the case that an amplification according to NAL-RP ( $\alpha=0$ ) is applied. The figure also contains the relevant hearing threshold. The same speech spectrum, but normalized to the hearing threshold, is shown in Fig. 2(c). The corresponding figures for the case of the simulated hearing loss are shown in Figs. 2(b) and 2(d). It can be seen that after amplification or attenuation, respectively, the (grey-shaded) area between the average target speech spectrum and the hearing threshold is

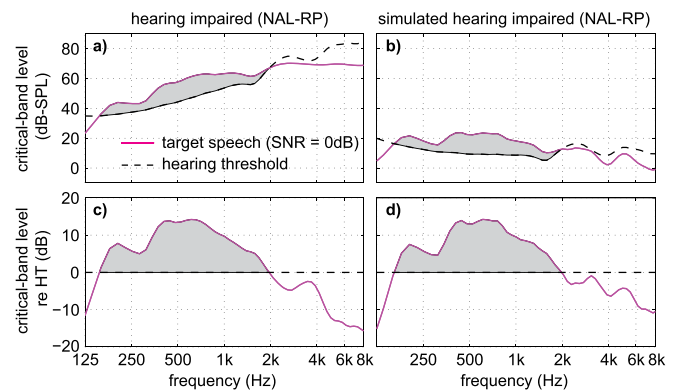


FIG. 2. (Color online) Illustration of the method applied to simulate a hearing loss for an example gain of NAL-RP and an SNR of 0 dB. Target speech levels for the average HI and simulated HI subject are shown in (a) and (b) together with their corresponding hearing thresholds. The same target speech spectra are shown in (c) and (d) but this time relative to their hearing thresholds. (a) is equal to both the NAL-RP condition and hearing threshold shown in Fig. 1(b) and is only included for comparison.

basically the same for the HI and the simulated HI case. The slightly rougher curve in the simulated HI case is due to the NH BP filterbank that was applied to derive the simulated HI speech spectra, which had narrower filters than in the HI case (see Sec. II C 1).

#### D. Procedure

Testing took place in a sound-attenuated booth. The LiSN-S was administered via HD215 Sennheiser headphones (Sennheiser, Wednebstel, Germany) attached to a personal computer using an in-house produced MATLAB version of the LiSN-S software (Glyde *et al.*, 2013a). The MATLAB (Mathworks, Natick, MA) version was selected for use in the experiment instead of the commercially available version to allow the experimenters to easily replace the original LiSN-S audio files with those that had the audibility filters applied. To ensure presentation levels prior to filtering would have been equivalent to that used in the commercial software, calibration was undertaken according to Cameron and Dillon (2007) using a GRAS RA0045 ear simulator.

An adaptive one-up one-down procedure was applied to measure speech reception threshold (SRT), the SNR at which a listener correctly understands 50% of all the words within a sentence. The starting SNR was +7 dB and participants were required to repeat the target sentences heard. Whole word scoring was used and no tense errors were accepted. If the participant repeated more than 50% of the words in the sentence correctly the SNR was decreased by 2 dB. If less than 50% of the words were repeated correctly the SNR was increased by 2 dB. If exactly 50% of the words were repeated correctly then the SNR remained the same. Testing started after at least five sentences were presented to the subject and the first upward reversal occurred. Testing in each condition concluded when either a minimum of 17 sentences were presented and the participant achieved a standard error of less than 1 dB or the maximum number of 30 sentences was reached. The participant's 50% SRT was then calculated as the average SNR at which each of the trials was presented during the testing phase. This procedure was the same as used in Glyde *et al.* (2013c).

#### E. Comparison samples

For the purposes of analysis, data from the current study were compared to results from two other samples: (1) 16 HI adults from Glyde *et al.* (2013c); and (2) the 96 NH adults from the LiSN-S normative data study (Cameron *et al.*, 2011). The 16 HI adults were chosen for comparison purposes as their hearing losses were similar to those of the HI participants in the current study (4FAHLs within 5 dB of 50 dB HL; see Table II for audiometric thresholds) and moreover, they received "standard" amplification according to NAL-RP. The normative data sample was selected to provide a reference point for LiSN-S performance in audiometrically NH adult listeners. Descriptive statistics regarding the age range for each group is provided in Table I.

### III. RESULTS

All statistical analyses were conducted using R (version 3.0.2), with the additional packages nlme (version 3.1–113) and multcomp (version 1.3–1).

The mean SRTs and 95% confidence intervals for each condition for simulated HI and HI individuals are displayed in Fig. 3. Figure 4 shows the mean SRM and 95% confidence intervals for both groups at the different audibility levels. In both the spatially separated (SV90) and co-located (SV0) conditions performance was consistently worse for the HI group at each audibility level. It also appears that for both groups increased stimulus levels resulted in better performance in the spatially separated condition but had little effect on the co-located condition. To assess whether this reflects significant changes in SRM, the spatial advantage results were analyzed further.

To account for the fact that the data contain more than one measurement for some subjects (see Table I), results from the HI group were analyzed by fitting a linear mixed-effects model (Hothorn *et al.*, 2008). This model had spatial advantage as the dependent variable, amplification level as a fixed effect, and a subject-specific intercept as the random effect. A significant effect of amplification level on spatial advantage was found [ $F_{(2,14)} = 13.9$ ,  $p < 0.001$ ], and all differences between pairs of levels were statistically significant after adjustment for multiple comparisons ( $p < 0.02$ ).

As for the HI group, a linear mixed-effects model was fitted to the simulated HI group, with the difference being that now there were four amplification levels instead of three. A significant effect of amplification on spatial advantage was found for the simulated HI groups [ $F_{(3,14)} = 34.9$ ,  $p < 0.001$ ], and all differences between pairs of levels were statistically significant after adjustment for multiple comparisons ( $p < 0.05$ ).

To determine whether SRM differed significantly between groups at any of the tested amplification levels (NAL-RP, NAL-RP + 25%, NAL-RP + 50%), a linear mixed-effects model was fitted to both groups but excluding the data for normal audibility. The dependent variable was spatial advantage, the fixed effects were group, amplification level and their interaction, and the random effect was a subject-specific intercept. Significant differences were found

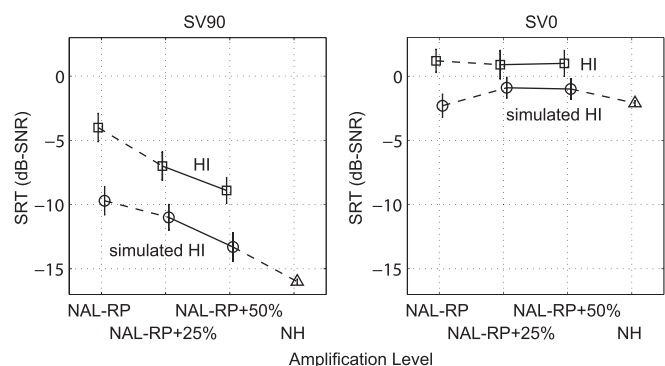


FIG. 3. Mean SRT  $\pm$  95% confidence intervals for SV90 and SV0 conditions by amplification level for both HI groups (squares) and simulated HI groups (circles). Results from comparison samples are connected by the dashed lines.

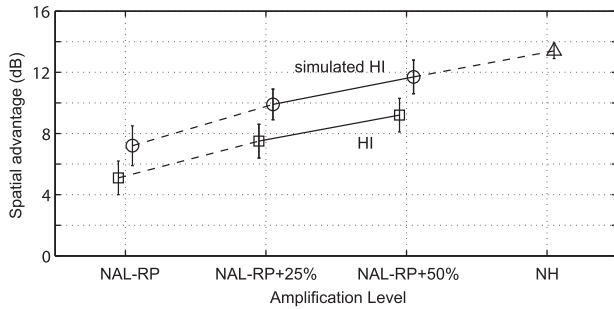


FIG. 4. Mean SRM  $\pm$  95% confidence intervals by amplification level for both the simulated HI (circles) and HI groups (squares). Results from the comparison samples are indicated by the dashed lines.

for the main effects group [ $F_{(1,58)} = 19.9, p < 0.001$ ] as well as amplification level [ $F_{(2,28)} = 33.1, p < 0.001$ ], but no interaction was found between group and amplification level [ $F_{(2,28)} = 0.1, p = 0.923$ ]. All differences between groups within each amplification level were statistically significant after adjustment for multiple comparisons ( $p < 0.02$ ).

## IV. DISCUSSION

### A. Summary of results

The aim of this study was to investigate whether reduced audibility may contribute to the reduced SRM that has been observed for HI listeners in symmetric speech-on-speech masking paradigms (e.g., Marrone *et al.*, 2008; Best *et al.*, 2012; Gallun *et al.*, 2013; Glyde *et al.*, 2013c). By reducing the audibility of speech materials presented to NH listeners as well as increasing the amplification provided to HI listeners, we were able to examine the effect of four different levels of audibility on SRM. For both the HI and simulated HI listeners, spatial advantage systematically improved as the provided amount of amplification increased. From the NAL-RP baseline, applying an extra gain of 25%/50% resulted in an improvement of 2.4/4.1 dB for the HI listeners and 2.7/4.5 dB for the simulated HI listeners. When spatial advantage is compared for the simulated HI group with NAL-RP to the comparison NH group with “full” audibility, the overall improvement was 6.2 dB. The observed improvements in spatial advantage were purely due to a decrease in SRTs achieved in the spatially separated condition. These results support the hypothesis that a reduction in audibility can limit the amount of SRM that an individual attains.

These results are consistent with the earlier work of Arbogast *et al.* (2002), who found that the sensation level of the stimuli significantly affected the amount of SRM achieved for a speech-on-speech masking condition. They reported that differences between their HI and NH groups lessened when discrepancies in sensational level were minimized. The findings of Kidd *et al.* (2010) are also consistent with the current findings. In that study, sentence materials presented in spatially separated, interfering speech were low-pass filtered at 1.5 kHz for NH listeners. Low-pass filtering would have resulted in a large reduction in, or complete loss of, audibility at high frequencies and, as in the present

study, SRM was found to be reduced compared to the broadband control condition.

### B. Possible consequences of reduced audibility

As laid out in Sec. I, reduced audibility has the potential to affect SRM in several ways, including a reduction in absolute target audibility, a loss of head shadow benefits when competing sounds are spatially separated, and reduced access to spatial cues for both binaural signal enhancement and localization.

Absolute target audibility can affect performance on speech-in-speech tasks by setting an upper limit on performance; if the target is not sufficiently above a listener’s hearing threshold (as illustrated in Fig. 5, right panel) then intelligibility will only reach a certain level no matter what cues (such as spatial separation) are provided. This issue is especially relevant for tests such as the LiSN-S that use an adaptive target level, where the target can reach very low levels. In the LiSN-S, low target levels are more likely to be reached in the spatially separated condition than in the co-located condition, and thus elevated hearing thresholds are more likely to reduce audibility in the spatially separated condition. This is a viable explanation for the pattern of results we observed, in which reduced audibility reduced the separated thresholds and thus the SRM.

To better understand the role of head shadow in the current experiment, we compared the spectrum levels of the maskers applied in the simulated HI condition with the spectrum level of threshold simulating noise. The idea is that when the level of the masker is above the level of the threshold simulating noise, then the masker is the dominant masker and its temporal, spectral, and spatial properties directly affect the audibility of the target. In such cases, the auditory system can apply better-ear glimpsing to improve speech intelligibility, assuming the target is also in the audible range (see above). In the opposite case, the threshold simulating noise is the dominant masker, which is here described by a diffuse noise without any direction-dependent properties,

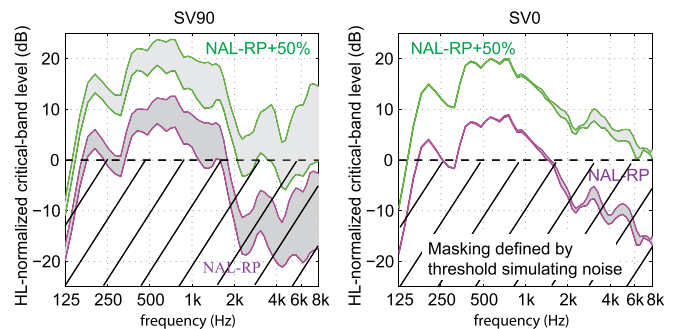


FIG. 5. (Color online) Masker spectrum levels at the output of a NH BP filterbank relative to the spectrum level of threshold simulating noise for two amplification levels. The spectrum level of the threshold simulating noise is given by the dashed lines. The larger of the two spectrum levels is considered the effective masker level. Shaded areas indicate the range of masker levels when averaged either over the ear with the lower or higher short-term level, i.e., indicating the potential effect of better-ear glimpsing. The average target spectrum at 0 dB-SNR is very similar to the spectrum of the co-located maskers shown in the right panel.

and hence there are no opportunities to make use of the head shadow.

The spectrum level of the threshold simulating noise was calculated using the procedures described in ANSI S3.5-1997 (ANSI, 1997). The reference internal noise spectrum level given in Table I of ANSI S-3.5-1997 was transformed to the equivalent level at the listener's eardrums by applying the HRTFs used in the LiSN-S to simulate target and masker signals that arrive from the front. The resulting spectrum levels were integrated in 1-ERB wide frequency bands. Similarly, the spectrum level of the maskers was calculated using a NH BP filterbank (see Sec. II C 2). To illustrate the potential effect of better-ear glimpsing, in each frequency channel, short-term RMS levels were calculated by averaging either the lower or higher of the levels at the two ears of the listener. The resulting better-ear and worse-ear spectrum levels of the maskers, normalized by the spectrum level of the threshold simulating noise, are shown in Fig. 5 for NAL-RP and NAL-RP + 50% amplification condition. The grey-shaded areas indicate the range from the better ear to the worse ear masker levels and thus, indicate the potential effect of better ear glimpsing. The left panel shows the case of the spatially separated condition and the right panel of the co-located condition. The spectrum level of the threshold simulating noise is indicated by the dashed lines, and due to the applied normalization is given by a straight horizontal line at a level of 0 dB. It should be noted that only long-term RMS signal levels are shown here, and that in particular for the highly fluctuating speech masker, the individual short-term RMS levels vary significantly around these long-term levels.

In the spatially separated condition (Fig. 5, left panel) it can be seen that when NAL-RP is applied, a major part of the masker that could be used for better-ear glimpsing is buried in the threshold simulating noise-floor and therefore, cannot be accessed by the auditory system to enhance speech intelligibility. When extra gain (sufficient to remove 50% of the deficit in audibility relative to normal hearing) is applied on top of NAL-RP, a significant part of the masker is lifted above the threshold simulating noise-floor and is available for better-ear glimpsing. In the co-located case (Fig. 5, right panel) better-ear glimpsing is negligible.

Besides better-ear glimpsing (or ILD-based processing), the auditory system can additionally utilize ITD cues to improve speech intelligibly in spatial noise, either by using an equalization-cancellation-like process (Durlach, 1963) to reduce EM, or by providing localization cues to spatially segregate the target from the maskers and reduce IM (Freyman *et al.*, 1999). ITD processing is much more effective at low frequencies (<1.5 kHz), where the signal's fine structure is analyzed by the auditory system, than at high frequencies, where the signal's envelope is analyzed (e.g., Blauert, 1997). Given that, in the current case, audibility was generally not an issue at low frequencies and amplification mainly affected audibility at high frequencies, it is expected that any advantage provided by ITDs was rather independent of amplification. Although we have focused on how audibility might affect the spatial cues that drive SRM, it is also worth considering the idea that audibility can affect the

relative amounts of EM and IM in a speech mixture, and hence indirectly influence the amount of SRM observed. For example, Arbogast *et al.* (2002) raised the idea that increasing the sensation level might increase the intelligibility of the competing sounds and hence the amount of IM experienced. Since, in some circumstances, a greater amount of masking means greater scope for unmasking, this could explain an increase in SRM. Following this argument, the co-located SRTs shown in Fig. 3 (left panel) should increase with increasing amplification, reflecting an increase in IM, and the spatially separated thresholds should remain relatively constant. Instead, we found that amplification mainly affected the separated thresholds, and had little effect on the co-located thresholds (in fact they tended to decrease – even though not significantly). Overall, the effects of amplification we observed are more consistent with an explanation based on target audibility, and/or the availability of spatial cues, rather than changes in IM. It is worth noting that reduced audibility caused by hearing loss clearly reduces SRM for stimuli containing *no* IM (such as when speech is masked by a steady noise, e.g., Dubno *et al.*, 2002).

### C. Remaining differences between NH and HI

Despite equating audibility between the simulated HI and HI participants in the present study, a significant difference in performance and SRM remained between the two groups at each of the amplification levels. The simulated HI participants performed approximately 1.5 dB better in the co-located condition and 4 dB better in the separated condition than the HI group of the current study for both the NAL-RP + 25% and NAL-RP + 50% conditions. Clearly the measured difference in the co-located condition, and part of the difference in the separated condition could be attributed to a number of factors that are not necessarily related to spatial separation. However, NH listeners who were tested with the filtered stimuli still achieved significantly greater SRM than HI individuals, achieving on average approximately 2.5 dB more benefit.

The difference in SRM between the groups may be attributable to the fact that the HI participants had a range of hearing thresholds, some worse and some better than the hearing loss we attempted to match with the stimulus presented to the simulated HI individuals (see Table II). It is possible that the listeners with worse thresholds brought down the mean performance in the HI group, by a greater degree than those with better thresholds brought it up, relative to those with the average loss. However, this explanation does not seem likely given the strong linear relationship between hearing thresholds and performance in each of the LiSN-S conditions found in Glyde *et al.* (2013c).

It is also possible that the 40 year average age difference between the HI groups and simulated HI groups may play a role. A number of researchers have argued that age may contribute to spatial processing deficits (e.g., Divenyi and Haupt, 1997; Divenyi *et al.*, 2005; Murphy *et al.*, 2006; Gallun *et al.*, 2013), although other studies have found no significant age effect for SRM (e.g., Gelfand *et al.*, 1988; Singh *et al.*, 2008; Cameron *et al.*, 2011; Füllgrabe *et al.*,

2015). It is worth noting that the effect of age, within an adult cohort, on SRM achieved in the LiSN-S paradigm was calculated using multiple regression analysis in the Glyde *et al.* (2013c) study and found to be insignificant ( $P = 0.10$ ). Nevertheless, those data showed that up to 1 dB change in spatial advantage might be expected between the mean age of the simulated HI sample and the mean age of the HI sample.

Another potential factor contributing to the SRM differences between the HI group and simulated HI participants is changes to frequency resolution within the frequency range where the target signal is audible. People with a sensorineural hearing loss are known to have wider auditory filters (de Boer and Bouwmeester, 1974) effectively resulting in a reduced number of independent auditory channels. This has the potential to affect performance in two ways. First, HI participants would have less opportunity to compare information across frequency channels and to identify some of the finer spectro-temporal differences between the target and interfering speech in both the spatially separated and co-located conditions, thus resulting in reduced overall performance. Second, the wider spread of excitation along the cochlea which is associated with wider auditory filters could affect the ability of a HI individual to access better-ear spectral glimpses or reduce the efficiency of the better ear glimpsing mechanism by providing fewer glimpses but with a broader bandwidth. For a similar cohort of HI individuals, Glyde *et al.* (2013a) predicted that reduced better-ear glimpsing efficiency due to decreased frequency resolution results in a 1 dB reduction in SRM.

Moreover, one should not rule out the potential contribution of ITDs to the group difference observed here. As mentioned earlier, ITDs are dominant in the low frequencies and would have been relatively unaffected by the filtering applied for the young NH listeners in this study. On the other hand, there is some evidence in the literature that listeners with hearing loss have reduced sensitivity to ITDs (Füllgrabe and Moore, 2014), which may reduce their ability to achieve SRM (e.g., Gallun *et al.*, 2014).

Regardless of the cause of this remaining 2.5 dB deficit in SRM, the finding that audibility can account for the majority of the observed reduction in SRM (5.8 dB of an initial 8.3 dB difference) is an important one. For researchers it suggests that audibility must be carefully accounted for if differences in SRM are to be fully understood. For hearing rehabilitation it suggests that if HI people could be provided with greater amplification, particularly in the mid to high frequencies (>2–3 kHz) as done here, their difficulty hearing in noise may be substantially reduced. This is by no means a straightforward task, since often the gain required cannot be provided by current hearing aids, and in any case listeners might not accept such high levels in their daily life because of their reduced dynamic range. Nonetheless, as shown by Moore *et al.* (2010), extended bandwidth nonlinear amplification can provide small improvements in speech intelligibility under conditions of spatial separation. This may be even further improved by novel interactive combinations of compression, directional amplification and adaptive noise reduction.

## V. CONCLUSIONS

This investigation determined that reduced audibility can account for a substantial proportion of the spatial processing deficits of HI listeners at least when SRM is measured using the applied speech-on-speech paradigm. Increasing amplification in HI individuals, on top of NAL-RP, resulted in increased SRM, while reducing audibility in NH listeners led to a reduction in SRM. However, after audibility was equalized the SRM measured in HI subjects was still consistently smaller than in NH listeners, suggesting that other factors besides audibility play a role.

## ACKNOWLEDGMENTS

The authors acknowledge the financial support of the HEARING CRC, established and supported under the Australian Government's Cooperative Research Centres Program, an Australian Government Initiative, and the financial support of the Commonwealth Department of Health and Ageing. V.B. was also partially supported by NIH-DIDCD grant DC04545.

- ANSI S3.5 (1997). revision of ANSI S3.5-1969 (R 1986): *Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).
- 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.
- Bamiou, D. E. (2007). "Measures of binaural interaction," in *Handbook of (Central) Auditory Processing Disorder, Auditory Neuroscience and Diagnosis*, edited by F. E. Musiek and G. D. Chermak (Plural Publishing, San Diego, CA), Vol. 1, pp. 257–286.
- Best, V., Mason, C. R., and Kidd, G., Jr. (2012). "The influence of non-spatial factors on measures of spatial release from masking," *J. Acoust. Soc. Am.* **131**, 3103–3110.
- Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization, Revised edition* (MIT Press, Cambridge, MA).
- Bronkhorst, A., and Plomp, R. (1988). "The effect of head-induced interaural time and level difference on speech intelligibility in noise," *J. Acoust. Soc. Am.* **83**, 1508–1516.
- Brungart, D. S., and Iyer, N. (2012). "Better-ear glimpsing efficiency with symmetrically-placed interfering talkers," *J. Acoust. Soc. Am.* **132**, 2545–2556.
- Cameron, S., and Dillon, H. (2007). "Development of the listening in spatialized noise-sentences test (LISN-S)," *Ear Hear.* **28**, 196–211.
- Cameron, S., Glyde, H., and Dillon, H. (2011). "Listening in spatialized noise-sentences test (LiSN-S): Normative and retest reliability data for adolescents and adults up to 60 years of age," *J. Am. Acad. Audiol.* **22**, 697–709.
- de Boer, E., and Bouwmeester, J. (1974). "Critical bands and sensorineural hearing loss," *Audiology* **13**, 236–259.
- Dillon, H. (2012). *Hearing Aids*, 2nd ed. (Boomerang Press, Sydney), pp. 290–297.
- Divenyi, P. L., and Haupt, K. M. (1997). "Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation," *Ear Hear.* **18**, 189–201.
- Divenyi, P. L., Stark, P. B., and Haupt, K. M. (2005). "Decline of speech understanding and auditory thresholds in the elderly," *J. Acoust. Soc. Am.* **118**, 1089–1100.
- 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.
- Durlach, N. I. (1963). "Equalization and cancellation theory of binaural masking-level differences," *J. Acoust. Soc. Am.* **35**, 1206–1218.
- 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**(4), 3578–3588.



- Füllgrabe, C., and Moore, B. C. J. (2014). "Effects of age and hearing loss on stream segregation based on interaural time differences," *J. Acoust. Soc. Am.* **136**(2), EL185–EL191.
- Füllgrabe, C., Moore, B. C. J., and Stone, M. A. (2015). "Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition," *Front. Aging Neurosci.* **6**, 347.
- Gallun, F. J., Diedesch, A. C., Kämpel, S. D., and Jakien, K. M. (2013). "Independent impacts of age and hearing loss on spatial release in a complex auditory environment," *Front. Neurosci.* **7**, 252.
- Gallun, F. J., McMillan, G. P., Molis, M. R., Kämpel, S. D., Dann, S. M., and Konrad-Martin, D. (2014). "Relating age and hearing loss to monaural, bilateral, and binaural temporal sensitivity," *Front. Neurosci.* **8**, 172.
- 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.
- Glyde, H., Buchholz, J., Dillon, H., Best, V., Hickson, L., and Cameron, S. (2013a). "The effect of better-ear glimpsing on spatial release from masking," *J. Acoust. Soc. Am.* **134**, 2937–2945.
- Glyde, H., Buchholz, J., Dillon, H., Cameron, S., and Hickson, L. (2013b). "The importance of interaural time differences and level differences in spatial release from masking," *J. Acoust. Soc. Am.* **134**, EL147–EL152.
- Glyde, H., Cameron, S., Dillon, H., Hickson, L., and Seeto, M. (2013c). "The effects of hearing impairment and aging on spatial processing," *Ear Hear.* **34**, 15–28.
- Hothorn, T., Bretz, F., and Westfall, P. (2008). "Simultaneous inference in general parametric models," *Biomet. J.* **50**, 346–363.
- Kidd, G., Jr., Mason, C. R., and Best, V. (2010). "Stimulus factors influencing spatial release from speech-on-speech masking," *J. Acoust. Soc. Am.* **128**, 1965–1978.
- Kidd, G., Jr., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. I. (2008). "Informational masking," in *Auditory Perception of Sound Sources*, Springer Handbook of Auditory Research, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer, New York), pp. 143–189.
- Marrone, N., Mason, C. R., and Kidd, G., Jr. (2008). "The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms," *J. Acoust. Soc. Am.* **124**, 3064–3075.
- Moore, B. C., Glasberg, B. R., and Stone, M. A. (2010). "Development of a new method for deriving initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF," *Int. J. Audiol.* **49**(3), 216–227.
- Murphy, D. R., Daneman, M., and Schneider, B. A. (2006). "Why do older adults have difficulty following conversations?," *Psych. Aging* **21**, 49–61.
- Nejime, Y., and Moore, B. C. J. (1997). "Simulation of the effect of threshold elevation and loudness recruitment combined with reduced frequency selectivity on the intelligibility of speech in noise," *J. Acoust. Soc. Am.* **102**, 603–615.
- Patterson, R. D., Holdsworth, J., Nimmo-Smith, I., and Rice, P. (1988). "Svos final report: The auditory filterbank," Tech. Rep. APU Rep. 2341.
- Singh, G., Pichora-Fuller, M. K., and Schneider, B. A. (2008). "The effect of age on auditory spatial attention in conditions of real and simulated spatial separation," *J. Acoust. Soc. Am.* **124**, 1294–1305.