

RESEARCH ARTICLE

# The Effects of Sensorineural Hearing Impairment on Asynchronous Glimpsing of Speech

Erol J. Ozmeral<sup>1\*</sup>, Emily Buss<sup>2</sup>, Joseph W. Hall, III<sup>2</sup>

**1** Department of Communication Sciences and Disorders, University of South Florida, Tampa, Florida, United States of America, **2** Department of Otolaryngology/Head and Neck Surgery, School of Medicine, University of North Carolina, Chapel Hill, North Carolina, United States of America

\* [eozermeral@usf.edu](mailto:eozermeral@usf.edu)



**OPEN ACCESS**

**Citation:** Ozmeral EJ, Buss E, Hall JW, III (2016) The Effects of Sensorineural Hearing Impairment on Asynchronous Glimpsing of Speech. PLoS ONE 11 (5): e0154920. doi:10.1371/journal.pone.0154920

**Editor:** Giuseppe Biagini, University of Modena and Reggio Emilia, ITALY

**Received:** December 14, 2015

**Accepted:** April 21, 2016

**Published:** May 4, 2016

**Copyright:** © 2016 Ozmeral et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** This work was supported by National Institutes of Health—National Institute on Deafness and Other Communication Disorders grants R01 DC000418 awarded to JWH and F31 DC012694 awarded to EJO. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

In a previous study with normal-hearing listeners, we evaluated consonant identification masked by two or more spectrally contiguous bands of noise, with asynchronous square-wave modulation applied to neighboring bands. Speech recognition thresholds were 5.1–8.5 dB better when neighboring bands were presented to different ears (dichotic) than when all bands were presented to one ear (monaural), depending on the spectral width of the frequency bands. This dichotic advantage was interpreted as reflecting masking release from peripheral spread of masking from neighboring frequency bands. The present study evaluated this effect in listeners with sensorineural hearing loss, a population more susceptible to spread of masking. Speech perception (vowel-consonant-vowel stimuli, as in /aBa/) was measured in the presence of fluctuating noise that was either modulated synchronously across frequency or asynchronously. Hearing-impaired listeners ( $n = 9$ ) and normal-hearing controls were tested at either the same intensity ( $n = 7$ ) or same sensation level ( $n = 8$ ). Hearing-impaired listeners had mild-to-moderate hearing loss and symmetrical, flat audiometric thresholds. While all groups of listeners performed better in the dichotic than monaural condition, this effect was smaller for the hearing-impaired (3.5 dB) and equivalent-sensation-level controls (3.3 dB) than controls tested at the same intensity (11.0 dB). The present study is consistent with the idea that dichotic presentation can improve speech-in-noise listening for hearing-impaired listeners, and may be enhanced when combined with amplification.

## Introduction

Recognizing speech in a spectro-temporally dynamic background relies, in part, on a listener’s ability to integrate speech cues from the time/frequency regions where the signal-to-noise ratio (SNR) is favorable [1–6]. The ability to use isolated segments of the speech stream, sometimes referred to as speech “glimpsing” [4], is enhanced in the presence of masker fluctuation relative to steady maskers. The benefit associated with this masker fluctuation has been called the

fluctuating masker benefit [7, 8] or masking release [9]. Hearing-impaired (HI) listeners are often shown to have less ability to benefit from dynamic changes in local SNR compared to normal-hearing (NH) listeners. Though the mechanisms responsible for this are not fully understood, a number of factors associated with hearing loss have been implicated, including reductions or deficiencies in: audibility [10], temporal resolution [10–12], frequency selectivity [13–17], temporal fine structure processing [18–20], across-frequency integration [21, 22], and effects related to SNR in the baseline condition [7]. The purpose of the present study was to specifically assess the roles that frequency selectivity and audibility have on HI listeners' limited ability to benefit from masker fluctuation for speech perception.

We used a unique masking release paradigm with temporally-modulated noise maskers that were either comodulated or uncomodulated across frequency [2]. So as not to be confused with a well-established psychoacoustic phenomenon, comodulated masking release (CMR [23]), we refer to the comodulated and uncomodulated maskers as synchronously- and asynchronously-modulated maskers, respectively. One motivation for testing these unique maskers is that asynchronously-modulated maskers are more ecologically relevant than synchronously-modulated maskers, as many natural listening environments contain multiple sound sources. However, the periodic modulation pattern and constant level of the asynchronously-modulated maskers does not capture the variability and unpredictability associated with many real-world sounds. Masking release associated with these modulated maskers has previously been reported for NH listeners [2, 24], but not for HI listeners, who are likely to show less masking release than NH listeners [7, 10–22].

Our previous work showed that glimpsing in spectro-temporally complex environments improves when negative effects of masking spread are removed [24]. For asynchronously-modulated maskers, in particular, potentially favorable SNRs in spectro-temporal regions of masker minima are influenced by neighboring spectro-temporal regions of masker maxima. The severity of masking spread is dependent on the frequency selectivity of the individual. This was previously assessed by measuring performance in the asynchronously-modulated masker presented either monaurally or dichotically. Dichotic presentation meant that alternating frequency regions were separated across the ears to avoid peripheral masking spread from proximal frequency regions [25–29]. Masking release in dichotic, asynchronously-modulated masker conditions was larger than in the monaural case, and this was interpreted as the direct result of removing negative effects of masking spread. Recent work by Stone and colleagues [8], however, suggests that dichotic presentation could have reduced masking by eliminating intermodulations resulting from an interaction between masker bands in the periphery. Nevertheless, the potential for a benefit in HI listeners from dichotic presentation has not been evaluated for these maskers, and results may provide further support for the role of frequency selectivity in masked speech perception.

A number of studies have indicated that frequency selectivity is often reduced in listeners with sensorineural hearing loss [30–33]. One manifestation of reduced frequency selectivity in hearing impairment is a greater effect of spread of masking [31, 34, 35], although not all studies have observed consistent differences between listeners with and without hearing loss ([36], for a review, see [37]). The HI listeners in the present study were expected to experience a robust benefit from dichotic presentation of stimuli in the asynchronously-modulated masker condition. Such a result would be consistent with an interpretation that poor frequency selectivity limits glimpsing in spectro-temporally complex backgrounds. We also considered the possibility that HI listeners would have a limited ability to integrate information across frequency. Some support for this possibility was reported by Healy and Bacon [21] and Healy and Carson [22], although this deficit has not been seen in all paradigms [5]. If such limitations are present for HI listeners, it is unclear whether these factors would limit the extent to which dichotic

presentation could help performance via reduced spread of masking. We included control conditions to evaluate spectro-temporal integration, allowing us to test whether HI listeners have comparable integration abilities to those seen in NH listeners [24]. Finally, the role of audibility was separately evaluated by including normal-hearing control listeners who received either equal intensity or equivalent sensation level as the HI test group.

The goal of the present study was to assess HI listeners' speech perception in conditions where masking spread could be alleviated through a dichotic manipulation. Results showed that HI listeners were susceptible to limits in masking release overall, believed to be an effect of poor audibility. Nonetheless, dichotic listening was demonstrated to be beneficial for HI listeners under some masking conditions, thereby leaving the possibility that frequency selectivity was a viable target for remediation. In addition, these results offer further insight into HI listeners' abilities to integrate speech glimpses across time and frequency [5, 22], and under certain scenarios, we believe dichotic presentation may lead to better speech perception for bilateral hearing-instrument users.

## Materials and Methods

### Ethics Statement

All testing followed the ethical guidelines provided by the National Institutes of Health of the United States of America. Subjects provided written informed consent prior to all test measures and were compensated for their participation. The study, including consent and compensation, was approved by the Institutional Review Board at the University of North Carolina at Chapel Hill.

### Listeners

Twenty-four native English-speaking adults were recruited from the local and surrounding communities. The HI group ( $n = 9$ ) received the same stimuli as the normal-hearing control group (NH,  $n = 7$ ), and additional data were collected from a secondary, normal-hearing group (NH<sub>SL</sub>,  $n = 8$ ) who received stimuli near the sensation level of the HI group. The NH and NH<sub>SL</sub> listeners had pure-tone thresholds of 20 dB HL or lower at octave frequencies from 0.25 to 8 kHz in each ear [38]. The HI listeners had bilateral mild-to-moderate sensorineural hearing loss of no more than 60 dB HL between 0.25 and 8 kHz. Thresholds were approximately symmetric ( $\leq 20$  dB difference between ears) and relatively flat ( $\leq 25$  dB difference between 500 and 4000 Hz in all but one ear). Flat hearing loss was desirable to ensure approximately comparable access to speech cues across the speech spectrum. Ages ranged from 21 to 68 years old and were roughly matched across NH and HI groups (NH group: mean 42.9 yrs  $\pm$  14.4 sd; HI group: mean 46.6 yrs  $\pm$  17.4 sd). Because the NH<sub>SL</sub> group was a secondary dataset, no attempt was made to match age with the two primary groups (mean 29.5 yrs  $\pm$  13.5 sd). Listeners over the age of 60 years (1 NH and 2 HI listeners) completed a cognitive assessment before the experiment (Montreal Cognitive Assessment; [39]). These older listeners were required to obtain a score of 26 or better for inclusion in the study, and all three met this criterion. Demographic information is reported in the left-most columns of Table 1, and Fig 1 presents the average audiograms for each group.

### Stimuli

Speech stimuli were identical to those used in an earlier study [24]. The speech material included five recordings each for 12 vowel-consonant-vowels ([b d f g k m n p s t v z] as in /aga/), spoken by an adult female speaker and recorded at a sampling rate of 44.1 kHz.

**Table 1. Individual age and audiometric data.**

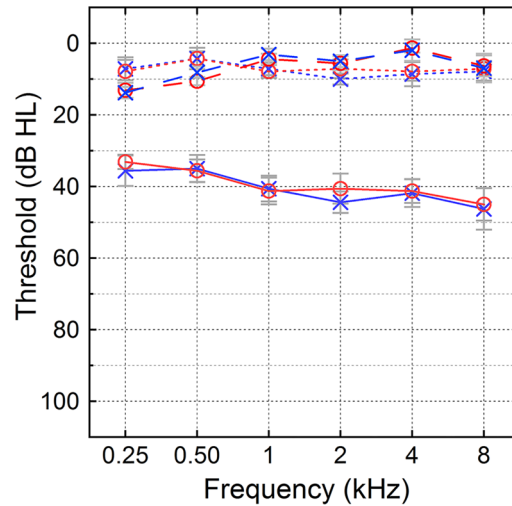
Group ID	Age (yrs)	PTA-L (dB)	PTA-R (dB)	VCV in Quiet (dB)
HI02	34.8	43.3	50.0	n/a
HI08	21.2	55.0	55.0	n/a
HI09	56.1	43.3	40.0	59.6
HI10	53.7	41.7	38.3	57.5
HI13	41.5	45.0	41.7	67.0
HI14	23.6	35.0	30.0	51.3
HI16	68.9	33.3	28.3	48.4
HI17	67.1	23.3	30.0	46.2
HI18	52.8	40.0	45.0	63.8
MEAN (sem)	46.6 (5.8)	40.0 (2.9)	39.8 (3.1)	56.2 (3.0)
Group ID	Age (yrs)	PTA-L (dB)	PTA-R (dB)	VCV in Quiet (dB)
NH01	18.4	13.3	15.0	28.0
NH03	44.8	6.7	6.7	33.6
NH04	38.1	3.3	-3.3	30.4
NH06	48.3	8.3	3.3	28.9
NH07	45.0	8.3	6.7	26.7
NH11	38.8	1.7	5.0	28.5
NH12	66.8	8.3	11.7	37.8
MEAN (sem)	42.9 (5.4)	7.1 (1.4)	6.4 (2.2)	30.6 (1.5)
Group ID	Age (yrs)	PTA-L (dB)	PTA-R (dB)	VCV in Quiet (dB)
NH <sub>SL</sub> 19	47.5	5.0	1.7	n/a
NH <sub>SL</sub> 20	54.3	10.0	5.0	n/a
NH <sub>SL</sub> 21	20.5	3.3	8.3	23.4
NH <sub>SL</sub> 22	25.6	5.0	8.3	25.0
NH <sub>SL</sub> 23	24.6	6.7	13.3	28.7
NH <sub>SL</sub> 24	21.5	6.7	5.0	27.9
NH <sub>SL</sub> 25	21.3	1.7	5.0	24.6
NH <sub>SL</sub> 26	20.5	5.0	8.3	31.1
MEAN (sem)	29.5 (4.8)	5.4 (0.9)	6.9 (1.2)	26.8 (1.2)

PTA-L, pure-tone average in left ear; PTA-R, pure-tone average in right ear; VCV, vowel-consonant-vowel threshold; HI, hearing-impaired; NH, normal-hearing; NH<sub>SL</sub>, normal-hearing with equivalent sensation level; sem, standard error of the mean

doi:10.1371/journal.pone.0154920.t001

Stimulus duration ranged from 528 to 664 ms, with a mean duration of 608 ms. Each token was normalized to equal root-mean-square level and filtered into 2, 4, 8, or 16 frequency bands using sixth-order Butterworth band-pass filters. For a given number of bands, filter bandwidths were equivalent in logarithmic units, with bands spanning 0.1 to 10 kHz.

As in previous studies using these methods, maskers were based on broadband pink noise samples which, by definition, contained equal energy per octave band. Each masker sample was generated digitally with duration equal to the longest possible speech token plus 300 ms (964 ms total duration). Speech stimuli began 150 ms after the onset of the noise masker. Masker modulation was either synchronous (Sync) or asynchronous (Async). Spectral representations of the modulated maskers are depicted in Fig 2, including asynchronously-modulated maskers with increasing numbers of filtered frequency bands. Sync maskers were modulated in the time-domain with a 10-Hz quasi-square wave with a random starting phase; 10-ms raised cosines were used to smooth level transitions and limit spectral splatter. To create



**Fig 1. Pure-tone audiometry for participants.** Audiograms for normal-hearing (NH;  $n = 7$ ; dotted lines), normal-hearing at equivalent sensation level (NH<sub>SL</sub>;  $n = 8$ ; dashed line), and hearing-impaired (HI;  $n = 9$ ; solid lines) listeners for both left (blue X's) and right (red circles) ears in dB HL. The HI listeners were screened to have roughly flat and symmetric mild-to-moderate hearing losses. Error bars represent one standard error of the mean.

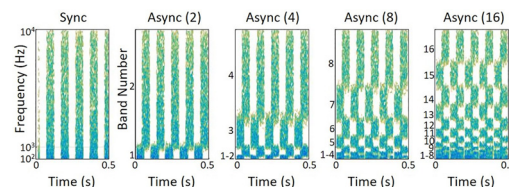
doi:10.1371/journal.pone.0154920.g001

Async maskers, the pink noise was filtered into 2, 4, 8, or 16 bands using sixth-order Butterworth band-pass filters. Then a 10-Hz quasi-square wave was applied to each noise band via multiplication. A single, randomly selected starting phase was chosen for the odd-numbered bands, and the inverse phase was used for the even-numbered bands. Bands were numbered by frequency region, beginning with the lowest frequency band. Before stimulus presentation, speech and noise signals were up-sampled to 48828 Hz to conform to hardware specifications (Tucker-Davis Technologies, Alachua, FL).

Either monaural (left [L] or right [R] ear only) or dichotic (D) stimuli were presented in a single block of trials. Monaural stimuli consisted of combined speech and noise signals. Dichotic stimuli included the odd-numbered bands of the combined speech and noise presented to the left ear, and even-numbered bands of the combined speech and noise presented to the right ear. In some cases, masker bands were presented to a single ear without the associated speech bands (see dichotic control conditions described below).

### Procedure and conditions

Procedures of the speech identification task were similar to those used in a previous study [24]. On each trial, speech tokens were randomly selected with replacement, and the task was to



**Fig 2. Spectrograms of modulated maskers used in the glimpsing task.** From left to right, the synchronously-modulated masker (Sync) and the asynchronously-modulated masker (Async) with 2, 4, 8, or 16 numbers of frequency bands. Modulation rates were set at 10 Hz, and initial phase of modulation was random.

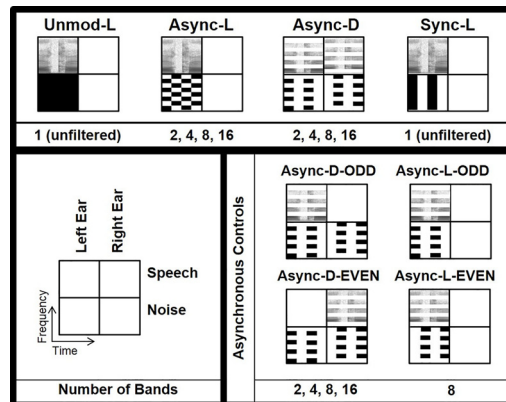
doi:10.1371/journal.pone.0154920.g002

indicate the consonant that was heard by selecting one of the 12 alternatives on the computer screen using a mouse. The speech recognition thresholds (SRTs) were measured using an adaptive up-down tracking algorithm using 4 dB steps, which estimated 50% correct identification [40]. The SRTs were determined based on the last 24 of 26 track reversals. This procedure was controlled using a custom Matlab (Mathworks, Inc., Natick, MA) script. Stimuli were presented through a pair of insert earphones (Etymotic ER-2, Elk Grove Village, IL), and listeners were seated in a single-wall, sound-treated booth. The first block of trials was a speech identification task with no masking. This served both to familiarize listeners with the task, and as a measure of in-quiet SRTs (results reported in Table 1). For the first two hearing-impaired and two normal-hearing listeners, in-quiet SRTs are not available due to an initial requirement of 100% speech identification accuracy at a comfortably loud presentation level. This initial requirement was subsequently replaced with an in-quiet threshold measure of speech reception.

For testing in quiet, thresholds were obtained by adjusting the level of the signal. Subsequent blocks of trials contained a masker. In these trials the SNR was adjusted adaptively, starting with an initial SNR of 10 dB. In the unmodulated noise condition, the level of the stimulus (signal plus masker) was fixed at 85 dB SPL for the HI and NH groups, and it was attenuated to 55 dB SPL for the NH<sub>SL</sub> group. To achieve the fixed level with varying SNR, two scalars were generated—one for the target, to produce the desired SNR, and the other for the summed target-plus-masker, to control the overall level. Both scalars were generated based on a bandpass filtered target (0.1–10 kHz) and a steady noise masker sample. This same procedure, generating a total level of 85 dB SPL, was the first stage for all other masked speech identification conditions. For the synchronous modulation, the noise was bandpass filtered and subsequently amplitude modulated. For the asynchronous monaural conditions, the noise was filtered and modulated before presentation. For the asynchronous dichotic conditions, the noise was modulated on a band-by-band basis prior to presentation. The full intensity was consequently reduced 3.2 dB by amplitude modulation in the Sync and Async conditions, and reduced further, in an ear-specific way, in the dichotic conditions. Trials were blocked by condition, and the order of conditions was quasi-randomly selected for each listener to avoid order effects. Each listener performed either three or four tracks for each condition. The fourth estimate was obtained if the first three thresholds were not all within 3 dB of each other. Overall testing time was roughly 5 h, typically spread out over five sessions on multiple days.

Fig 3 illustrates the key features of the 28 total conditions described in the remainder of this paragraph. In the baseline conditions, unmodulated noise was presented monaurally to either the left or right ear (Unmod-L and Unmod-R). The Sync condition was presented monaurally to each ear as well (Sync-L and Sync-R). For each Async monaural and dichotic condition (Async-L, Async-R and Async-D, respectively), stimuli were processed into 2, 4, 8, or 16 bands for a total of twelve Async test conditions. Additionally, there were two control conditions for the Async-D conditions. The first set of control conditions presented the Async-D masker (with 2, 4, 8, or 16 bands) but included only half of the speech bands: in Async-D-EVEN, the even-numbered speech bands were presented to the right ear, and in Async-D-ODD, the odd-numbered speech bands were presented to the left ear. These control conditions were intended to reveal whether performance in the Async-D conditions could be accounted for solely by either the even or odd speech bands alone. By including the masker in both ears but speech in only one ear, we were also able to test the possibility that contralateral maskers could affect performance. Two additional control conditions were included to assess masking in the Async-D condition from a single ear. In the Async-L-ODD and Async-L-EVEN conditions, only the odd-numbered or even-numbered frequency bands were presented to the left ear, respectively (the right-ear conditions were also tested but are not depicted in Fig 3). These conditions were





**Fig 3. Schematics of monaural and dichotic masking conditions.** Primary conditions are represented on the top row, and control conditions are shown below. Only left-ear (L) schematics are visually depicted for monaural conditions, but right (R) ear conditions were also tested. As the legend indicates, each condition is represented as a 2-by-2 box in which the left and right columns represent stimulation of the left and right ears, respectively, and the top and bottom rows represent the speech and noise stimuli, respectively. In each box, frequency from 0.1 to 10 kHz is represented vertically, and a time span of 200 ms is represented horizontally. Speech is represented via spectrogram, and noise is represented in black. Amplitude modulation is performed at a rate of 10 Hz, and frequency bands are equally spaced on a logarithmic scale. The order of the primary conditions in the top row is an indication of the expected ranking in thresholds, with the worst performance starting on the left, with the Unmod-L and Unmod-R conditions, and the best performance on the right, with the Sync-L and Sync-R conditions. The numbers of bands tested per condition are given below each condition schematic.

doi:10.1371/journal.pone.0154920.g003

only run using 8 band-pass filters (i.e., 4 bands per ear). All conditions were tested for the HI and NH groups, whereas the NH<sub>SL</sub> group was tested only on conditions with 4 or 8 bands.

### Data analysis

Although we tested monaural conditions in both left and right ears, the performance obtained from each of the two ears was very closely matched in both the NH or HI groups. For example, the absolute value of the difference in SRTs for the Unmod-L and Unmod-R conditions was (on average) 0.8 dB for NH listeners and 1.2 dB for HI listeners. Given the similarity across ears, data were analyzed after taking the average of the left and right ear SRTs. The resulting values are identified with an ‘M’ (e.g., Unmod-M) to indicate monaural presentation. Moreover, reporting of data in the control conditions was limited to the best-case performance. For instance, on a subject-by-subject basis, the better threshold in either the Async-D-ODD or Async-D-EVEN was the only dichotic control threshold used to assess the performance on control conditions. The better of the two dichotic control conditions is reported as Control-D, and the better of the monaural control conditions is reported as Control-M. We used the lower (better) of the two control thresholds to evaluate performance in the primary Async conditions because it would provide the most conservative measure of integration when all bands were available. Additionally, this simplification mitigates effects related to subtle asymmetries in hearing between ears.

Data in each test condition were submitted to a Shapiro-Wilk test of normality. Significant values were obtained in only two cases: the 16-band, monaural Async condition for the NH group ( $p = 0.001$ ) and the 8-band, dichotic Async condition for the NH<sub>SL</sub> group ( $p = 0.01$ ). It was decided to conduct parametric analyses despite evidence of non-normality in these two conditions due to simplicity and ease of interpretation.

## Results

### Hearing-impaired listener thresholds

Mean SRTs for the HI listeners are presented in Table 2 (top) for all primary conditions and the better of the control conditions. To measure the ability to glimpse speech in a fluctuating masker, data were analyzed in terms of masking release, quantified as the difference in SRT between a condition with modulated noise and the Unmod-M condition. Fig 4 (left panel) shows masking release (in dB) for the average of the monaural Async conditions (Async-M), the dichotic condition (Async-D), the average of the Sync conditions (Sync-M), and better of the Async-D control conditions (Control-D), expressed relative to the SRT for the Unmod-M reference value. Error bars show one standard error of the mean, and symbols indicate the masker condition, as defined in the legend. The shaded region at the bottom of the figure indicates the range of values that can be accounted for by the fact that modulation reduces the overall masker level by 3.2 dB. The masking release for HI listeners was greatest for Sync-M (average of 8.0 dB) and for Async-D (ranging from 6.0 to 9.4 dB), but it was consistently smaller for Async-M (ranging from 2.4 to 5.9 dB). Masking release was evaluated with single-sample one-tailed t-tests, with a reference of 3.2 dB (the reduction in masker level associated with modulation). Masking release was greater than 3.2 dB for all four Async-D conditions and for the 2-band Async-M condition ( $p < 0.05$ ), but not for the other Async-M conditions ( $p \geq 0.421$ ). A two-way repeated-measures ANOVA was performed to compare performance

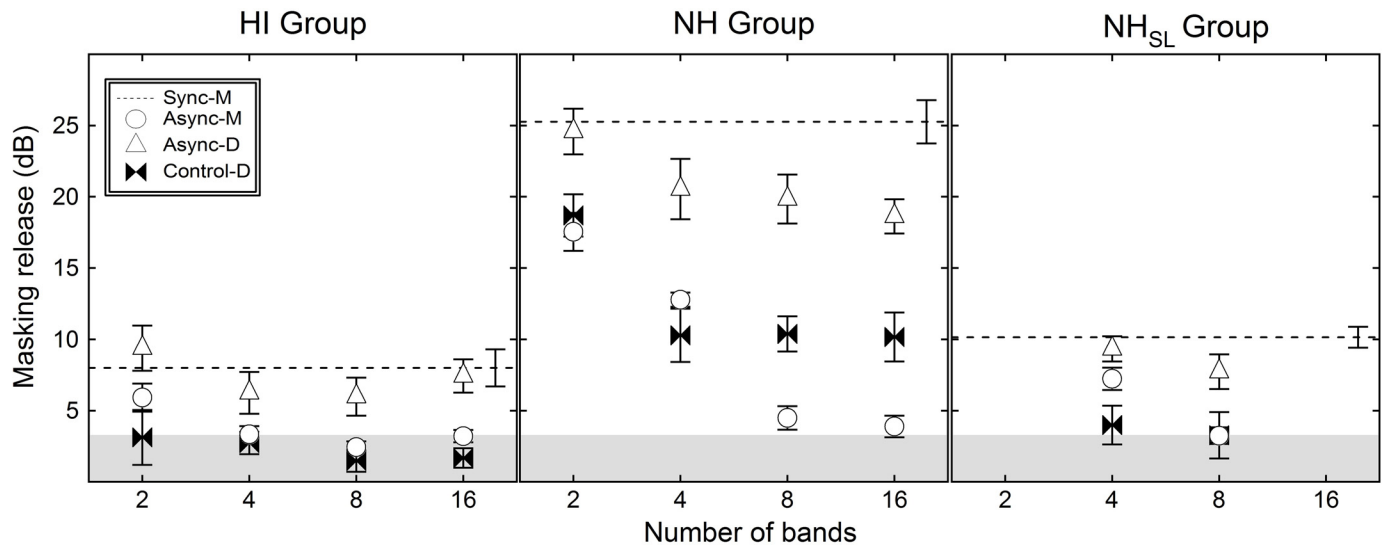
**Table 2. Mean speech recognition thresholds (in dB SNR) for each primary test condition and the better of the control conditions.**

HI		Number of Bands			
		2	4	8	16
Unmod-M	0.3 (0.5)				
Sync-M	-7.7 (1.2)				
Async-M		-5.6 (0.9)	-3.0 (0.5)	-2.1 (0.5)	-2.9 (0.4)
Async-D		-9.1 (1.3)	-6.0 (1.4)	-5.7 (1.2)	-7.1 (1.0)
Control-D		-2.8 (1.8)	-2.4 (0.7)	-1.5 (0.6)	-1.4 (0.8)
Control-M				-5.4 (1.4)	
NH		Number of Bands			
		2	4	8	16
Unmod-M	-0.3 (0.6)				
Sync-M	-25.5 (1.5)				
Async-M		-17.8 (1.3)	-13.0 (0.6)	-4.7 (1.2)	-4.1 (0.8)
Async-D		-24.9 (1.7)	-20.8 (1.8)	-20.1 (1.7)	-18.9 (1.0)
Control-D		-19.0 (1.4)	-10.6 (1.8)	-10.7 (1.1)	-10.4 (1.7)
Control-M				-14.7 (1.5)	
NH <sub>SL</sub>		Number of Bands			
		2	4	8	16
Unmod-M	-2.2 (0.6)				
Sync-M	-12.4 (1.1)				
Async-M			-9.6 (1.1)	-5.3 (1.1)	
Async-D			-11.6 (1.3)	-10.0 (1.6)	
Control-D			-4.0 (1.4)	-3.3 (1.6)	

HI, hearing-impaired; NH, normal-hearing; NH<sub>SL</sub>, normal-hearing with equivalent sensation level; Unmod-M, monaural unmodulated masker; Sync-M, monaural synchronously-modulated masker; Async-M, monaural asynchronously-modulated masker; Async-D, dichotic asynchronously-modulated masker; Control-M, monaural controls; Control-D, dichotic controls

doi:10.1371/journal.pone.0154920.t002





**Fig 4. Masking release relative to unmodulated noise.** Mean masking release is plotted for modulated noise conditions relative to the unmodulated condition for hearing-impaired (HI; left panel) normal controls with stimulus presentation at either the same intensity (NH; middle panel) or same sensation level (NH<sub>SL</sub>; right panel). The difference in mean thresholds relative to the Unmod-M condition at 2, 4, 8, and 16 bands are plotted for the monaural asynchronous condition (Async-M; circles), the dichotic asynchronous condition (Async-D; triangles), the better of the dichotic control conditions (Control-D; bowties), and the mean of the synchronous conditions (Sync-M; dashed line). Error bars indicate standard error of the mean (n = 9 for HI group; n = 7 for NH group; n = 8 for NH<sub>SL</sub> group). The shaded region at the bottom of each panel indicates the range over which masking release might be accounted for by reductions in masker level associated with amplitude modulation (3.2 dB).

doi:10.1371/journal.pone.0154920.g004

in the Async-D and Async-M conditions, with two levels of condition and four levels of band number. This analysis yielded a main effect of condition ( $F_{1,8} = 10.2, p = 0.013$ ), a main effect of the number of bands ( $F_{3, 24} = 15.7, p < 0.001$ ), but no interaction ( $F_{3, 24} = 0.43, p = 0.73$ ). Simple main effects testing was performed to compare masking release in the Async-D to the Async-M conditions; in all cases more masking release was observed in the dichotic than the monaural presentation condition ( $p < 0.05$ , with Bonferroni correction), as was the case for NH listeners in the previous study [24].

### Normal-hearing listeners—Equal Intensity

Mean SRTs for NH listeners tested at 85 dB SPL are presented in Table 2 (middle) for all primary test conditions and the better dichotic control conditions. Fig 4 (middle panel) shows the mean masking release (in dB) of the NH group for the Async-M, Async-D, Sync-M, and Control-D, measured relative to the SRT for the Unmod-M reference. Error bars show one standard error of the mean, and symbols indicate the masker condition, as defined in the legend. The masking release for NH listeners was greatest for Sync-M (average of 25.3 dB), intermediate for Async-D (ranging from 18.6 to 24.6 dB), and least for Async-M (ranging from 3.9 to 17.5 dB), with some values in this last condition being consistent with a reduction in overall masker level (in contrast to glimpsing). Masking release was significantly greater than 3.2 dB for all modulated masker conditions ( $p < 0.001$ ) in all Async-M and Async-D conditions *except* the Async-M-8 condition ( $p = 0.090$ ) and the Async-M-16 condition ( $p = 0.211$ ). A two-way repeated-measures ANOVA was performed to compare performance in Async-D and Async-M, with two levels of condition and four levels of band number. This analysis yielded a main effect of condition ( $F_{1,6} = 108.8, p < 0.001$ ), a main effect of the number of bands ( $F_{3,18} = 64.5, p < 0.001$ ), and an interaction ( $F_{3,18} = 15.1, p < 0.001$ ). The interaction is explained by the greater separation between masking release observed in the different conditions as the

band number increased. Simple main effects testing was performed to compare masking release in the Async-D to the Async-M conditions; in all cases better performance was observed in the dichotic than the monaural presentation condition ( $p \leq 0.005$ , with Bonferroni correction).

### Normal-hearing listeners—Equal sensation level

Mean SRTs for NH<sub>SL</sub> listeners tested at 55 dB SPL are presented in [Table 2](#) (bottom). [Fig 4](#) (right panel) shows the mean masking release (in dB) of the NH<sub>SL</sub> group for the Async-M, Async-D, Sync-M, and Control-D, measured relative to the SRT for the Unmod-M reference. Error bars show one standard error of the mean, and symbols indicate the masker condition, as defined in the legend. The masking release for NH<sub>SL</sub> listeners was greatest for Sync-M (average of 10.2 dB), intermediate for Async-D (4 bands: 9.3 dB; 8 bands: 7.7 dB), and least for Async-M (4 bands: 7.4 dB; 8 bands: 3.1 dB). Masking release was significantly greater than 3.2 dB for all modulated masker conditions ( $p < 0.01$ ) *except* the Async-M-8 condition ( $p = 0.85$ ). A two-way repeated-measures ANOVA was performed to compare performance in Async-D and Async-M, with two levels of condition and two levels of band number. This analysis yielded a main effect of condition ( $F_{1,7} = 34.0$ ,  $p = 0.001$ ), a main effect of the number of bands ( $F_{1,7} = 186.8$ ,  $p < 0.001$ ), and an interaction ( $F_{1,7} = 7.9$ ,  $p < 0.05$ ). As before with the NH group, the interaction is explained by the greater separation between masking release observed in the different conditions as the band number increased from 4 to 8 bands. Moreover, post-hoc testing showed greater masking release in the dichotic than the monaural presentation for both 4 and 8 bands ( $p < 0.005$ , with Bonferroni correction).

### Between-group analyses

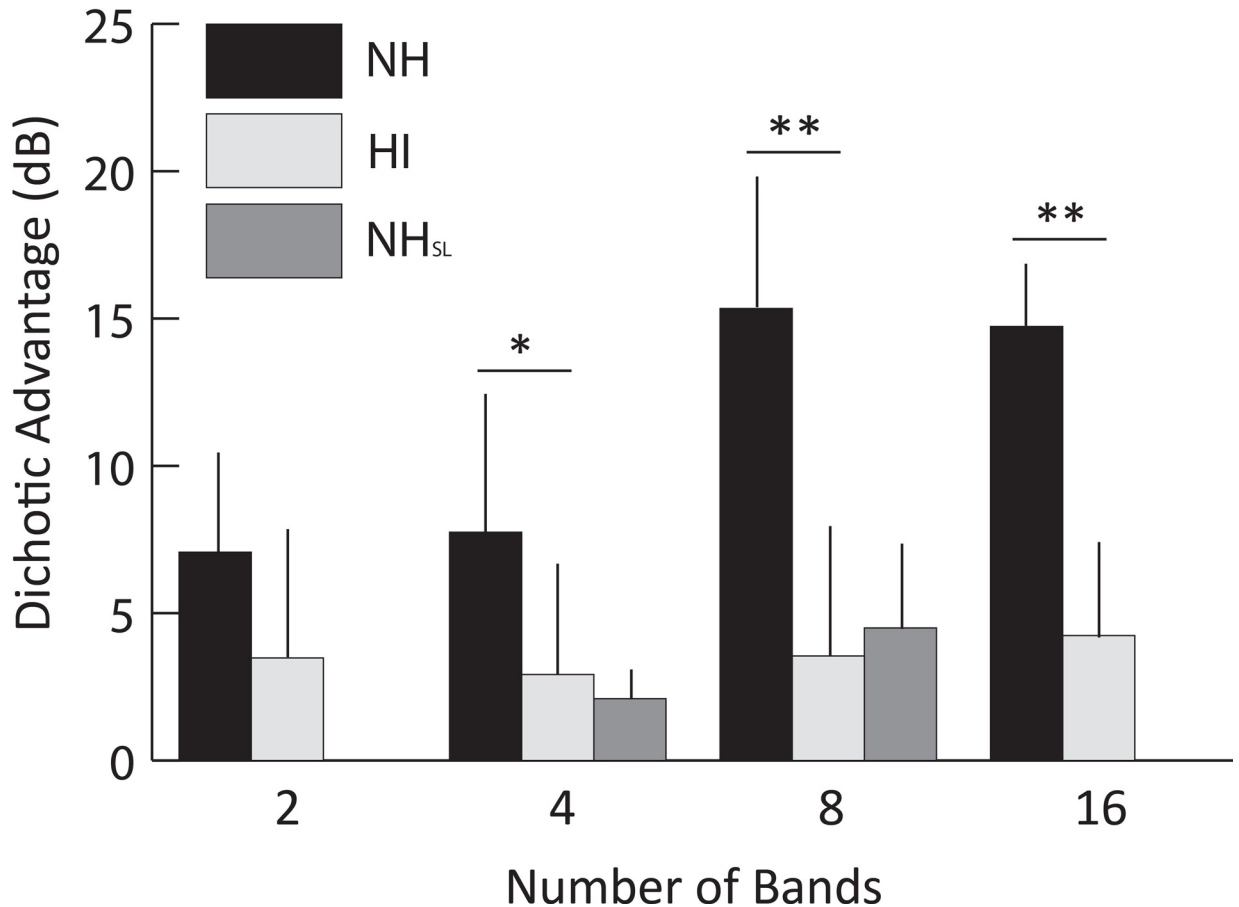
**Comparisons at same intensity.** The SRTs in the Unmod-M case were submitted to a one-way ANOVA. This analysis showed no significant difference between the NH and HI listeners ( $F_{1,14} = 0.53$ ,  $p = 0.48$ ), which indicated that at an overall presentation level of 85 dB SPL, hearing impairment did not reliably affect speech recognition in steady noise. It is evident from [Fig 4](#), however, that NH listeners had greater masking release in most modulated-noise conditions compared to the HI group. A one-way ANOVA for Sync-M masking release indicated that the difference was significant ( $F_{1,14} = 90.55$ ,  $p < 0.001$ ). With respect to the Async noise conditions, masking release data were submitted to a three-way ANOVA with two levels of presentation type (dichotic and monaural), four levels of number of bands (2, 4, 8, and 16), and two levels of listener group (NH and HI). This analysis showed significant main effects of presentation type ( $F_{1,14} = 87.4$ ,  $p < 0.001$ ), number of bands ( $F_{3,42} = 77.9$ ,  $p < 0.001$ ), and listener group ( $F_{1,14} = 83.3$ ,  $p < 0.001$ ). There were also significant interactions between condition and band number ( $F_{3,42} = 12.4$ ,  $p < 0.001$ ), between condition and group ( $F_{1,14} = 23.6$ ,  $p < 0.001$ ), and between band number and group ( $F_{3,42} = 27.7$ ,  $p < 0.001$ ). Lastly, the three-way interaction was significant ( $F_{3,42} = 8.8$ ,  $p < 0.001$ ). Because the three-way interaction was significant, the other interactions and significant main effects should be interpreted with caution. The significant three-way interaction is best explained by a large and relatively constant difference between groups for the Async-D conditions, yet in the Async-M conditions, the difference between NH and HI listeners was large for the 2 and 4 band numbers and vanishingly small by 16 bands. Because modulation caused a drop in overall intensity of the masker by 3.2 dB, the Async-M thresholds probably did not reflect glimpsing for 4, 8, and 16 bands in the HI listeners, or for 8 and 16 bands for the NH listeners. From [Fig 4](#), we can also see that while NH listeners tended to have less masking release in both noise conditions as the number of bands increased, HI listeners showed relatively consistent and low masking release for all numbers of bands. The absence of an effect of band number in the HI data may be influenced by the

compressed range of thresholds, including a relatively small peak masking release for HI listeners in the Sync-M condition. This final point can be addressed by comparing data for HI and NH listeners at similar sensation levels.

**Comparisons at equivalent sensation level.** For the NH<sub>SL</sub> group, average threshold in the baseline, Unmod-M condition was -2.2 dB SNR; that value was significantly lower than Unmod-M threshold for the HI group ( $F_{1,15} = 11.7$ ,  $p < 0.005$ ). The NH<sub>SL</sub> and HI masking release data were submitted to a three-way ANOVA with two levels of presentation type (dichotic and monaural), two levels of number of bands (4 and 8), and two levels of listener group (NH<sub>SL</sub> and HI). This analysis showed significant main effects of condition ( $F_{1,15} = 19.4$ ,  $p = 0.001$ ), number of bands ( $F_{1,15} = 35.5$ ,  $p < 0.001$ ), and listener group ( $F_{1,15} = 4.63$ ,  $p < 0.05$ ). There were also significant interactions between condition and band number ( $F_{1,15} = 5.0$ ,  $p < 0.05$ ) and between band number and group ( $F_{1,15} = 15.3$ ,  $p = 0.001$ ). There was no significant interaction between group and condition ( $p = 0.97$ ), nor was there a significant three-way interaction ( $p = 0.21$ ). The interaction between condition and band number appears to be explained by a greater rate of reduction in masking release from 4 to 8 bands in the monaural condition relative to the dichotic condition. The interaction between band number and group is apparent in the steeper decline in masking release from 4 to 8 bands for the NH<sub>SL</sub> group than the HI group.

Masking release data from each condition, including the controls and Sync-M conditions, were submitted to one-way ANOVAs comparing the NH<sub>SL</sub> and HI groups. The only significant difference between groups was in the 4-band Async-M condition ( $F_{1,15} = 16.8$ ,  $p = 0.001$ ). Inspection of each panel of Fig 4 shows that as the number of bands increases, masking release in the Async-M condition approaches floor at some point. Even when matched for sensation level, performance at floor occurs at a lower number of bands (4) in the HI group than in the NH<sub>SL</sub> group, where floor performance is not reached until 8 bands.

**Comparisons of dichotic advantage.** The differences in masking release between Async-D and Async-M conditions—referred to as dichotic advantage—are presented in Fig 5 for NH (black bars) and HI (shaded bars) listeners. The dichotic advantage was between 7.1 and 15.3 dB for the NH group, and between 2.9 and 4.2 dB for the HI group. A two-way ANOVA with two levels of group and four levels of number of bands resulted in a main effect of group ( $F_{1,14} = 23.5$ ,  $p < 0.001$ ), a main effect of number of bands ( $F_{3,42} = 12.4$ ,  $p < 0.001$ ), and a significant interaction ( $F_{3,42} = 8.8$ ,  $p < 0.001$ ). Simple main effects indicate that NH listeners had greater dichotic advantage than those with hearing loss for 4 bands ( $p < 0.05$ ), 8 bands ( $p < 0.001$ ) and 16 bands ( $p < 0.001$ ), but not for 2 bands ( $p = 0.096$ ). This is due to the fact that the dichotic advantage increased with number of bands for the NH group, but did not increase as much (if at all) for the HI group. Again, it is possible that the magnitudes of differences across conditions in the HI listeners are limited due to their smaller maximum masking release in the synchronous modulation condition. This was possible to assess from the NH<sub>SL</sub> group data, in which masking release in the Sync-M condition was more comparable between the listener groups. Dichotic advantage was analyzed for the NH<sub>SL</sub> and HI groups by submitting data to a two-way ANOVA with factors of group and number of bands. Although there was a significant main effect of number of bands ( $F_{1,15} = 5.0$ ,  $p < 0.05$ ), there was no significant effect of group or interaction between group and number of bands. The lack of an interaction is particularly interesting in light of the fact that HI listeners performed more poorly than the NH<sub>SL</sub> group in the 4-band, Async-M condition. This reveals that although HI listeners had more difficulty glimpsing speech in the presence of neighboring noise bands, the degree of benefit they received from dichotic listening was comparable to the NH listeners tested at a comparable sensation levels.



**Fig 5. The effect of hearing loss on dichotic benefit.** Dichotic benefit (i.e., the difference between Async-D and Async-M conditions) for hearing-impaired (HI) and normal hearing controls (NH and NH<sub>SL</sub>). Error bars indicate one standard deviation. Whereas the dichotic advantage increased significantly for NH listeners as band number increased, HI and NH<sub>SL</sub> listeners received similar benefit at all band numbers (NH<sub>SL</sub> group was only tested at 4 and 8 bands). Significant differences between groups are indicated by an asterisk ( $p < 0.05$ ) or two asterisks ( $p < 0.001$ ).

doi:10.1371/journal.pone.0154920.g005

### Evaluation of spectro-temporal integration

Control measures taken in the study were useful in assessing the possibility that a listener was attending just to a subset of bands—either the even or the odd bands—in the Async conditions, thereby not actually integrating across frequency *and* time. Performance in the Async-D conditions was uniformly better than either Async-D-ODD or Async-D-EVEN control conditions for both groups. The difference in SRTs between the Async-D condition and the better of the two control conditions ranged from 5.9 to 10.2 dB for NH subjects and from 3.5 to 6.3 dB for HI subjects, depending on the number of bands. This result supports the interpretation that speech perception in the Async-D condition was not based on cues present in either subset of bands presented to a single ear, but rather relied on cues distributed across ears.

Recall that in the Async-D-EVEN and Async-D-ODD conditions, the noise-only ear received bands of noise that were modulated out-of-phase relative to the masker modulation in the ear presented with the speech signal. We compared the Control-D and Control-M measures at 8 bands to assess the effect of including asynchronously modulated masker bands contralateral to the speech-plus-noise stimulus. Masking release in the monaural control conditions was 3.5 dB greater than in the dichotic control conditions for NH listeners; this

difference was 4.0 dB for HI listeners. From these results, it appears that having a modulated masker in non-overlapping frequency regions in one ear can mask speech in the other ear. The across-ear masking indicated by the comparison between the monaural and dichotic control conditions may help explain why listeners do not obtain as much release from masking in the Async-D conditions as they do in the Sync condition. That is, although the effects of masking spread have been reduced in the Async-D condition, there appears to be some other factor limiting performance in the dichotic condition. One possibility is that a contralateral masker affects perceptual weighting based upon masker modulation phase and listening in the dips [41]. Because the masker modulation phase in the contralateral ear was antiphase with respect to the modulation in the speech ear, it is possible that the contralateral masker resulted in “miscoinciding” that blunted the benefit of improved SNR associated with masker dips in the speech ear.

## Discussion

### Factors contributing to reduced masking release in HI listeners

**Effect of audibility.** Overall, HI listeners in the present study had less masking release than the NH group. Data from the NH<sub>SL</sub> group provided some indication that much of the difference between HI and NH listeners could be attributed to audibility. Previous studies have shown that HI listeners benefit less from masker amplitude modulation than NH listeners when stimuli are presented at equal levels [17, 42–45], especially for single syllable stimuli [46]. Although some of these results can be explained by reduced audibility in the masker dips for low-level speech cues, it has been suggested that other contributors are poor temporal resolution [10, 46] or poor frequency resolution [14, 17]. In the equivalent-intensity configuration, overall levels were fixed at 85 dB SPL; however, speech signals could fall well below that in some conditions, and low-level speech cues can be very important for identification [47]. For example, the target speech is approximately 60 dB SPL at -25 dB SNR, the approximate best SRT obtained in NH listeners; while a 60-dB-SPL target would be detectable in quiet for all HI listeners tested, some of the low-level cues would likely be inaudible, limiting those listeners’ ability to glimpse speech in the modulated maskers.

Another factor to consider when comparing masking release between listening groups is the threshold difference in the reference condition. Generally, NH listeners achieve larger masking release when the baseline SNR is more negative [48], which is related to the performance-intensity function of speech perception in noise [7]. The performance-intensity function indicates how much change in speech recognition is associated with a change in level—at medium levels, small changes in level will lead to large performance differences, whereas at low and high levels, small changes in level do not affect performance as much. Whereas numerous studies have shown that HI listeners are less able to benefit from the introduction of masker fluctuation compared to NH listeners [10, 17, 42], Bernstein and Grant [7] note that these particular studies were undermined by a confound between group differences in the baseline SNR. In the present study, baseline SNRs were not found to be significantly different between NH and HI groups, so this issue was less of a concern. The lack of baseline SRT differences between the NH and HI listeners is somewhat surprising because most previous studies have found elevated masked SRTs in listeners with sensorineural hearing loss [7, 49, 50]. For elevated presentation levels such as ours, however, NH listeners have been shown to be more closely aligned with HI listeners in masked-speech performance [47]. This possibility is supported by the data from the NH<sub>SL</sub> group which did differ significantly from the HI group in baseline SRT (-2.2 dB SNR and 0.3 dB SNR, respectively). However, the two groups did not consistently differ in their benefit from masker modulation, so this is a clear divergence from previous reports. Nevertheless, we considered two additional factors that may have contributed to the lack of a baseline SRT

difference between HI and NH groups. First, steeply sloping hearing loss has a greater detrimental effect on consonant recognition than gradually sloping or flat losses [51, 52]. The HI listeners in the present study had relatively flat audiograms, which would be associated with relatively modest effects of hearing loss on masked SRTs. Second, the use of pink noise in the present study could have affected performance. Whereas speech-spectrum noise falls off at approximately 8 dB/octave [53], pink noise falls off at 3 dB/octave. Because pink noise is relatively less effective at masking low- than high-frequency speech features, this masker could increase listeners' reliance on low-frequency cues. For the HI listeners in the present study, the hearing loss was relatively mild at low frequencies, which could have played a role in their good performance relative to NH listeners. This possibility is undermined, however, by the finding that flat mild/moderate hearing loss reduces performance for word recognition in pink noise [50].

**Effect of frequency selectivity.** Performance by all listeners was better for the asynchronous modulation conditions when stimuli were presented dichotically rather than monaurally. On average, this dichotic advantage was 7–15 dB for the NH listeners, which was even larger than previously seen (roughly 5–8 dB benefit in [24]). One methodological difference between studies that might account for this difference is presentation level. The current procedure presented stimuli at an overall level of 85 dB SPL (before modulation or separation of bands), whereas the previous experiment fixed the target level at 55 dB SPL and varied the masker level to estimate threshold. Therefore it should not be as surprising, considering that masking release has been shown previously to be smaller at lower intensities for both the synchronously- [10] and asynchronously-modulated maskers (e.g., Experiment 2 in [24]). Consequently, there was no difference between HI and NH<sub>SL</sub> groups in dichotic advantage. At similar sensation levels, HI listeners were able to integrate across spectro-temporal glimpses as well as normal-hearing controls when negative effects of peripheral spread of masking were removed. The lone difference between HI and NH<sub>SL</sub> groups was in the 4-band, monaural Async condition. Whereas each group performed equally poorly in the 8-band, monaural Async condition (i.e., no better than the 3.2 dB level difference accounted for by the modulation), the HI listeners also performed poorly in the 4-band case. This result was another clear indication that for these spectrally wide glimpsing regions, audibility alone could not account for the extent of the poor performance by the HI group. Indeed, the poor performance of the HI listeners in the 4-band monaural Async condition was probably due to reduced frequency selectivity. This interpretation is consistent with the finding that performance improved when the stimuli were presented dichotically.

## Possible clinical applications of dichotic listening

The results of this study have important implications for hearing aid design. In quiet settings, most aided HI listeners with mild-to-moderate sensorineural hearing loss have minor difficulty following a conversation. However, the same listeners often complain that it is difficult to follow speech in noisy environments. Traditional hearing aids with advanced noise-reduction processing [54] have been largely ineffective in improving speech understanding in noise. One of the obvious factors contributing to this phenomenon is that amplification has the negative effect of adding gain to all incoming sounds, including the unwanted noise. Therefore, supplementary strategies, like dichotic presentation, could be utilized to limit the influence unwanted noise has on speech perception.

There have been previous attempts to use dichotic presentation to improve speech identification in hearing-aid or cochlear implants users [25–28, 55], and the current study provides additional support for this approach. However, there could be unintended consequences of



removing crucial binaural spatial cues, such as interaural time or level differences [29], so further study in spatially diverse settings is still needed. Along with the support of amplification, dichotic presentation should be considered as an appropriate strategy for improving speech-in-noise performance.

## Conclusions

Previous studies using monaural Async maskers [2, 24] showed a decrease in the ability to benefit from masker modulation with increasing number of bands. One possible reason for this was increased spread of masking effects as the band number increased. Ozmeral *et al.* [24] aimed to reduce the possible deleterious effects of spread of masking by presenting neighboring spectral bands to separate ears [25]. The result was 5-to-8-dB better SRTs across all band conditions in the Async-D condition relative to the Async-M condition. The current study replicated NH listener data reported by Ozmeral *et al.* [24], and added the HI group to determine whether listeners with sensorineural hearing loss could also benefit from dichotic listening in the presence of an Async masker. Because HI listeners tend to have poorer-than-normal frequency selectivity [13–17], it was hypothesized that masking release would be greatly reduced or absent in a monaural asynchronous masker, but that dichotic presentation could facilitate masking release.

At equal presentation levels, listeners with sensorineural hearing loss had less masking release than age-matched normal-hearing listeners for speech presented in synchronously- and asynchronously-modulated noise. Testing at similar sensation levels between groups, however, indicated that masking release was closely associated with overall audibility. Importantly, the dichotic listening benefit was equivalent between HI and NH<sub>SL</sub> groups. These results are consistent with an interpretation that the reduced masking release shown by the HI listeners in monaural asynchronously-modulated noise is due to a combination of reduced audibility and poor frequency selectivity, and that amplification along with dichotic stimulation may provide the best outcomes for speech in spectro-temporally complex noise.

## Acknowledgments

A subset of these data were presented at the 2013 mid-winter meeting of the Association for Research in Otolaryngology. We thank Peter Gordon, Joseph Hopfinger, and Mark Hollins for consultations throughout this study, and special thanks to the PlosOne academic editor and two anonymous reviewers for their helpful comments on the manuscript.

## Author Contributions

Conceived and designed the experiments: EJO EB JWH. Performed the experiments: EJO EB. Analyzed the data: EJO EB JWH. Wrote the paper: EJO EB JWH.

## References

1. Miller GA, Licklider JCR. The intelligibility of interrupted speech. *J Acoust Soc Am.* 1950; 22(2):167–3. doi: <http://dx.doi.org/10.1121/1.1906584>
2. Howard-Jones PA, Rosen S. Unmodulated glimpsing in "checkerboard" noise. *J Acoust Soc Am.* 1993; 93(5):2915–22. Epub 1993/05/01. PMID: [8315155](https://pubmed.ncbi.nlm.nih.gov/8315155/).
3. Buss E, Hall JW, 3rd, Grose JH. Spectral integration of synchronous and asynchronous cues to consonant identification. *J Acoust Soc Am.* 2004; 115(5 Pt 1):2278–85. Epub 2004/05/14. PMID: [15139639](https://pubmed.ncbi.nlm.nih.gov/15139639/).
4. Cooke M. A glimpsing model of speech perception in noise. *J Acoust Soc Am.* 2006; 119(3):1562–73. Epub 2006/04/06. PMID: [16583901](https://pubmed.ncbi.nlm.nih.gov/16583901/).
5. Hall JW 3rd, Buss E, Grose JH. Spectral integration of speech bands in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am.* 2008; 124(2):1105–15. PMID: [18681600](https://pubmed.ncbi.nlm.nih.gov/18681600/). doi: [10.1121/1.2940582](https://doi.org/10.1121/1.2940582)

6. Assmann PF, Summerfield AQ. The perception of speech under adverse conditions. In: Greenberg S, Ainsworth WA, Popper AN, Fay RR, editors. *Speech Processing in the Auditory System*. 14. New York: Springer-Verlag, Inc.; 2004. p. 231–308.
7. Bernstein JG, Grant KW. Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2009; 125(5):3358–72. Epub 2009/05/12. doi: [10.1121/1.31110132](https://doi.org/10.1121/1.31110132) PMID: [19425676](https://pubmed.ncbi.nlm.nih.gov/19425676/).
8. Stone MA, Moore BCJ. On the near non-existence of "pure" energetic masking release for speech. *J Acoust Soc Am*. 2014; 135(4):1967–77. doi: [10.1121/1.4868392](https://doi.org/10.1121/1.4868392) PMID: [25234995](https://pubmed.ncbi.nlm.nih.gov/25234995/)
9. Fullgrabe C, Berthommier F, Lorenzi C. Masking release for consonant features in temporally fluctuating background noise. *Hear Res*. 2006; 211(1–2):74–84. Epub 2005/11/18. doi: [S0378-5955\(05\)00299-6](https://doi.org/S0378-5955(05)00299-6) [pii] doi: [10.1016/j.heares.2005.09.001](https://doi.org/10.1016/j.heares.2005.09.001) PMID: [16289579](https://pubmed.ncbi.nlm.nih.gov/16289579/).
10. George EL, Festen JM, Houtgast T. Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2006; 120(4):2295–311. Epub 2006/10/31. PMID: [17069325](https://pubmed.ncbi.nlm.nih.gov/17069325/).
11. Dubno JR, Horwitz AR, Ahlstrom JB. Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *J Acoust Soc Am*. 2002; 111(6):2897–907. PMID: [12083223](https://pubmed.ncbi.nlm.nih.gov/12083223/).
12. Dubno JR, Horwitz AR, Ahlstrom JB. Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. *J Acoust Soc Am*. 2003; 113(4 Pt 1):2084–94. Epub 2003/04/22. PMID: [12703719](https://pubmed.ncbi.nlm.nih.gov/12703719/).
13. Rosen SM, Fourcin AJ. Frequency selectivity and the perception of speech. *Frequency Selectivity in Hearing* 1986:373–488.
14. Baer T, Moore BCJ. Effects of Spectral Smearing on the Intelligibility of Sentences in Noise. *J Acoust Soc Am*. 1993; 94(3):1229–41. PMID: [ISI:A1993LX11800005](https://pubmed.ncbi.nlm.nih.gov/151A1993LX11800005/).
15. ter Keurs M, Festen JM, Plomp R. Effect of spectral envelope smearing on speech reception. II. *J Acoust Soc Am*. 1993; 93(3):1547–52. Epub 1993/03/01. PMID: [8473608](https://pubmed.ncbi.nlm.nih.gov/8473608/).
16. Baer T, Moore BC. Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J Acoust Soc Am*. 1994; 95(4):2277–80. PMID: [8201124](https://pubmed.ncbi.nlm.nih.gov/8201124/).
17. Peters RW, Moore BC, Baer T. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J Acoust Soc Am*. 1998; 103(1):577–87. PMID: [9440343](https://pubmed.ncbi.nlm.nih.gov/9440343/).
18. Qin MK, Oxenham AJ. Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. *J Acoust Soc Am*. 2006; 119(4):2417–26. PMID: [16642854](https://pubmed.ncbi.nlm.nih.gov/16642854/).
19. Gnansia D, Pean V, Meyer B, Lorenzi C. Effects of spectral smearing and temporal fine structure degradation on speech masking release. *J Acoust Soc Am*. 2009; 125(6):4023–33. Epub 2009/06/11. doi: [10.1121/1.3126344](https://doi.org/10.1121/1.3126344) PMID: [19507983](https://pubmed.ncbi.nlm.nih.gov/19507983/).
20. Hopkins K, Moore BC. The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *J Acoust Soc Am*. 2009; 125(1):442–6. Epub 2009/01/29. doi: [10.1121/1.3037233](https://doi.org/10.1121/1.3037233) PMID: [19173429](https://pubmed.ncbi.nlm.nih.gov/19173429/).
21. Healy EW, Bacon SP. Across-frequency comparison of temporal speech information by listeners with normal and impaired hearing. *J Speech Lang Hear Res*. 2002; 45(6):1262–75. PMID: [12546492](https://pubmed.ncbi.nlm.nih.gov/12546492/).
22. Healy EW, Carson KA. Influence of broad auditory tuning on across-frequency integration of speech patterns. *J Speech Lang Hear Res*. 2010; 53(5):1087–95. doi: [10.1044/1092-4388\(2010\)09-0185](https://doi.org/10.1044/1092-4388(2010)09-0185) PMID: [20689025](https://pubmed.ncbi.nlm.nih.gov/20689025/); PubMed Central PMCID: [PMC2954411](https://pubmed.ncbi.nlm.nih.gov/PMC2954411/).
23. Hall JW, Tyler RS, Fernandes MA. Factors influencing the masking level difference in cochlear hearing-impaired and normal-hearing listeners. *J Speech Hear Res*. 1984; 27(1):145–54. Epub 1984/03/01. PMID: [6717000](https://pubmed.ncbi.nlm.nih.gov/6717000/).
24. Ozmeral EJ, Buss E, Hall JW, 3rd. Asynchronous glimpsing of speech: spread of masking and task set-size. *J Acoust Soc Am*. 2012; 132(2):1152–64. Epub 2012/08/17. doi: [10.1121/1.4730976](https://doi.org/10.1121/1.4730976) PMID: [22894234](https://pubmed.ncbi.nlm.nih.gov/22894234/); PubMed Central PMCID: [PMC3427370](https://pubmed.ncbi.nlm.nih.gov/PMC3427370/).
25. Lunner T, Arlinger S, Hellgren J. 8-channel digital filter bank for hearing aid use: preliminary results in monaural, diotic and dichotic modes. *Scand Audiol Suppl*. 1993; 38:75–81. PMID: [8153568](https://pubmed.ncbi.nlm.nih.gov/8153568/).
26. Kulkarni PN, Pandey PC, Jangamashetti DS. Binaural dichotic presentation to reduce the effects of spectral masking in moderate bilateral sensorineural hearing loss. *Int J Audiol*. 2012; 51(4):334–44. Epub 2011/12/29. doi: [10.3109/14992027.2011.642012](https://doi.org/10.3109/14992027.2011.642012) PMID: [22201526](https://pubmed.ncbi.nlm.nih.gov/22201526/).
27. Loizou PC, Mani A, Dorman MF. Dichotic speech recognition in noise using reduced spectral cues. *J Acoust Soc Am*. 2003; 114(1):475–83. Epub 2003/07/26. PMID: [12880058](https://pubmed.ncbi.nlm.nih.gov/12880058/).

28. Zhou N, Pfungst BE. Psychophysically based site selection coupled with dichotic stimulation improves speech recognition in noise with bilateral cochlear implants. *J Acoust Soc Am*. 2012; 132(2):994–1008. Epub 2012/08/17. doi: [10.1121/1.4730907](https://doi.org/10.1121/1.4730907) PMID: [22894220](https://pubmed.ncbi.nlm.nih.gov/22894220/); PubMed Central PMCID: PMC3427365.
29. Aronoff JM, Amano-Kusumoto A, Itoh M, Soli SD. The effect of interleaved filters on normal hearing listeners' perception of binaural cues. *Ear Hear*. 2014; 35(6):708–10. doi: [10.1097/AUD.000000000000060](https://doi.org/10.1097/AUD.000000000000060) PMID: [25090455](https://pubmed.ncbi.nlm.nih.gov/25090455/); PubMed Central PMCID: PMC4243471.
30. Zwicker E, Schorn K. Psychoacoustical tuning curves in audiology. *Audiology*. 1978; 17(2):120–40. PMID: [WOS:A1978EU51100003](https://pubmed.ncbi.nlm.nih.gov/WOS:A1978EU51100003/).
31. Florentine M, Buus S, Scharf B, Zwicker E. Frequency selectivity in normally-hearing and hearing-impaired observers. *J Speech Hear Res*. 1980; 23(3):646–69. PMID: [WOS:A1980KM68100015](https://pubmed.ncbi.nlm.nih.gov/WOS:A1980KM68100015/).
32. Dubno JR, Dirks DD. Auditory Filter Characteristics and Consonant Recognition for Hearing-Impaired Listeners. *J Acoust Soc Am*. 1989; 85(4):1666–75. PMID: [ISI:A1989U235000032](https://pubmed.ncbi.nlm.nih.gov/ISI:A1989U235000032/).
33. Leek MR, Summers V. Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise. *J Acoust Soc Am*. 1993; 94(6):3127–37. doi: [10.1121/1.407218](https://doi.org/10.1121/1.407218) PMID: [WOS:A1993MM12600010](https://pubmed.ncbi.nlm.nih.gov/WOS:A1993MM12600010/).
34. Leshowitz B. Relationship of tonal masking to speech intelligibility in noise for listeners with sensorineural hearing damage. *J Acoust Soc Am*. 1977; 62:S93–S. doi: [10.1121/1.2016459](https://doi.org/10.1121/1.2016459) PMID: [WOS:A1977EA29000449](https://pubmed.ncbi.nlm.nih.gov/WOS:A1977EA29000449/).
35. Gagné JP. Excess masking among listeners with a sensorineural hearing loss. *J Acoust Soc Am*. 1988; 83(6):2311–21. PMID: [WOS:A1988P021000037](https://pubmed.ncbi.nlm.nih.gov/WOS:A1988P021000037/).
36. Martin ES, Pickett JM. Sensorineural hearing loss and upward spread of masking. *J Speech Hear Res*. 1970; 13(2):426–37. PMID: [WOS:A1970Y307600022](https://pubmed.ncbi.nlm.nih.gov/WOS:A1970Y307600022/).
37. Tyler RS. Frequency resolution in hearing-impaired listeners. *Frequency Selectivity in Hearing*. 1986:309–71.
38. ANSI. ANSI S3.6–2010. American National Standard Specification for Audiometers. New York: American National Standards Institute; 2010.
39. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *J Am Acad Ger Soc*. 2005; 53:695–9.
40. Levitt H. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*. 1971; 49(2):467–77. Epub 1971/02/01. PMID: [5541744](https://pubmed.ncbi.nlm.nih.gov/5541744/).
41. Buus S. Release from masking caused by envelope fluctuations. *J Acoust Soc Am*. 1985; 78(6):1958–65. Epub 1985/12/01. PMID: [4078172](https://pubmed.ncbi.nlm.nih.gov/4078172/).
42. Festen JM, Plomp R. Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *J Acoust Soc Am*. 1990; 88:1725–36. PMID: [2262629](https://pubmed.ncbi.nlm.nih.gov/2262629/)
43. Bronkhorst AW, Plomp R. Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *J Acoust Soc Am*. 1992; 92(6):3132–9. Epub 1992/12/01. PMID: [1474228](https://pubmed.ncbi.nlm.nih.gov/1474228/).
44. Gustafsson HA, Arlinger SD. Masking of speech by amplitude-modulated noise. *J Acoust Soc Am*. 1994; 95(1):518–29. Epub 1994/01/01. PMID: [8120263](https://pubmed.ncbi.nlm.nih.gov/8120263/).
45. Snell KB, Mapes FM, Hickman ED, Frisina DR. Word recognition in competing babble and the effects of age, temporal processing, and absolute sensitivity. *J Acoust Soc Am*. 2002; 112(2):720–7. PMID: [12186051](https://pubmed.ncbi.nlm.nih.gov/12186051/)
46. Jin SH, Nelson PB. Speech perception in gated noise: the effects of temporal resolution. *J Acoust Soc Am*. 2006; 119(5 Pt 1):3097–108. Epub 2006/05/20. PMID: [16708964](https://pubmed.ncbi.nlm.nih.gov/16708964/).
47. Studebaker GA, Sherbecoe RL, McDaniel DM, Gwaltney CA. Monosyllabic word recognition at higher-than-normal speech and noise levels. *J Acoust Soc Am*. 1999; 105(4):2431–44. PMID: [10212424](https://pubmed.ncbi.nlm.nih.gov/10212424/).
48. Oxenham AJ, Simonson AM. Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. *J Acoust Soc Am*. 2009; 125(1):457–68. doi: [10.1121/1.3021299](https://doi.org/10.1121/1.3021299) PMID: [19173431](https://pubmed.ncbi.nlm.nih.gov/19173431/); PubMed Central PMCID: PMC2677273.
49. Smoorenburg GF. Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. *J Acoust Soc Am*. 1992; 91(1):421–37. PMID: [1737889](https://pubmed.ncbi.nlm.nih.gov/1737889/).
50. Pekkarinen E, Salmivalli A, Suonpaa J. Effect of noise on word discrimination by subjects with impaired hearing, compared with those with normal hearing. *Scand Audiol*. 1990; 19(1):31–6. PMID: [2336538](https://pubmed.ncbi.nlm.nih.gov/2336538/).
51. Dubno JR, Dirks DD, Langhofer LR. Evaluation of hearing-impaired listeners using a Nonsense-syllable Test. II. Syllable recognition and consonant confusion patterns. *J Speech Hear Res*. 1982; 25(1):141–8. PMID: [7087417](https://pubmed.ncbi.nlm.nih.gov/7087417/).

52. Gordon-Salant S. Phoneme feature perception in noise by normal-hearing and hearing-impaired subjects. *J Speech Hear Res.* 1985; 28(1):87–95. PMID: [3982002](#).
53. Byrne D, Dillon H, Tran K, Arlinger S, Wilbraham K, Cox R, et al. An International Comparison of Long-Term Average Speech Spectra. *J Acoust Soc Am.* 1994; 96(4):2108–20. doi: [10.1121/1.410152](#) PMID: [WOS:A1994PL70800015](#).
54. Hu Y, Loizou PC. A comparative intelligibility study of single-microphone noise reduction algorithms. *J Acoust Soc Am.* 2007; 122(3):1777. Epub 2007/10/12. doi: [10.1121/1.2766778](#) PMID: [17927437](#).
55. Lyregaard PE. Frequency selectivity and speech intelligibility in noise. *Scand Audiol Suppl.* 1982; 15:113–22. PMID: [6955919](#).