Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention

Lengshi Dai[a], Virginia Best[b], and Barbara G. Shinn-Cunningham[a,1]

[a]Biomedical Engineering, Boston University, Boston, MA 02215; and [b]Speech, Language, & Hearing Sciences, Boston University, Boston, MA 02215

Edited by Eric I. Knudsen, Stanford University School of Medicine, Stanford, CA, and approved March 1, 2018 (received for review December 6, 2017)

Listeners with sensorineural hearing loss often have trouble understanding speech amid other voices. While poor spatial hearing is often implicated, direct evidence is weak; moreover, studies suggest that reduced audibility and degraded spectrotemporal coding may explain such problems. We hypothesized that poor spatial acuity leads to difficulty deploying selective attention, which normally filters out distracting sounds. In listeners with normal hearing, selective attention causes changes in the neural responses evoked by competing sounds, which can be used to quantify the effectiveness of attentional control. Here, we used behavior and electroencephalography to explore whether control of selective auditory attention is degraded in hearing-impaired (HI) listeners. Normal-hearing (NH) and HI listeners identified a simple melody presented simultaneously with two competing melodies, each simulated from different lateral angles. We quantified performance and attentional modulation of cortical responses evoked by these competing streams. Compared with NH listeners, HI listeners had poorer sensitivity to spatial cues, performed more poorly on the selective attention task, and showed less robust attentional modulation of cortical responses. Moreover, across NH and HI individuals, these measures were correlated. While both groups showed cortical suppression of distracting streams, this modulation was weaker in HI listeners, especially when attending to a target at midline, surrounded by competing streams. These findings suggest that hearing loss interferes with the ability to filter out sound sources based on location, contributing to communication difficulties in social situations. These findings also have implications for technologies aiming to use neural signals to guide hearing aid processing.

binaural hearing | electroencephalography | hearing impairment

Hearing-impaired (HI) listeners generally have poor spatial acuity (see review in ref. 1) and also show deficits in understanding speech from one direction when there is competing speech from other directions (2–4). It seems logical that these problems are related, yet studies looking for relationships between spatial acuity and speech-in-noise performance have provided equivocal results. While some find such relationships (e.g., ref. 5), others do not (6). In addition, some recent studies suggest that other nonspatial problems explain poor performance of HI listeners on speech-in-noise tasks (7). We hypothesized that poor spatial acuity is directly related to problems when understanding speech amid spatially separated competitors because it interferes with spatial selective auditory attention. To show that the problem is with selective attention, rather than due to the fact that understanding a degraded speech signal may be problematic for HI listeners, our main task was a spatial selective auditory attention task using nonspeech melodies.

Selective auditory attention depends on the formation of auditory objects (8–11); without properly parsing a scene, the brain cannot suppress responses to unattended sounds. Object formation is driven by local spectrotemporal structures as well as the continuity over time of higher-order features like pitch, timbre, and spatial location (12–14). Selective attention to auditory objects is influenced by both top-down control and bottom-up salience (14). Auditory attention operates as a form of sensory gain control, enhancing the representation of an attended object and suppressing the representation of ignored objects (15–17). A number of recent studies have demonstrated that the modulation of auditory-evoked responses by attention can be decoded from both electroencephalography (EEG) and magnetoencephalography (18–21).

In addition to having higher sound detection thresholds, listeners with sensorineural hearing impairment have reduced temporal and spectral acuity compared with normal-hearing (NH) listeners. We (and others) have argued in the past that this loss of acuity may degrade the representation of local structure and the features upon which object formation is based, and thus might weaken or slow down object formation, interrupting selective attention (9). Although this is a compelling idea, there is little evidence directly showing that HI listeners suffer from impairments related to poor auditory object segregation, such as difficulties suppressing competing sound.

Here, we used behavior and EEG to compare the efficacy of selective auditory attention in listeners with normal hearing and with bilateral, sloping sensorineural hearing loss. We used a previously published task in which listeners identify a simple melody presented simultaneously with two competing melodies at different spatial locations, while concurrently measuring EEG (15, 22). By design, early cortical responses to notes in each stream are separable in time, allowing us to quantify how strongly attention modulates the neural representation of a

Significance

Listeners with hearing loss have trouble communicating in many social settings. In such situations, listeners with normal hearing suppress the neural representation of competing sounds, a process known as auditory selective attention. We compared hearing-impaired (HI) and normal-hearing (NH) listeners on a spatial selective attention task while measuring neural responses using electroencephalography. Compared with NH listeners, HI listeners have poorer sensitivity to spatial cues, perform more poorly on the selective attention task, and show weaker neural suppression of competing sounds. Moreover, these different measures are correlated both for HI and for NH listeners. These results suggest that poor spatial acuity produces problems with selective attention. These findings have implications both for basic science and for development of next-generation hearing aids.

Author contributions: V.B. and B.G.S.-C. designed research; L.D. performed research; L.D. and B.G.S.-C. analyzed data; and L.D., V.B., and B.G.S.-C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: shinn@bu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1721226115/-/DCSupplemental.

Published online March 19, 2018.

sound mixture. We also tested sensitivity to interaural time differences (ITDs), a spatial cue that allows listeners to focus attention in this task. We found that, at the group level, HI listeners are less sensitive to ITDs, worse on the selective auditory attention task, and less able to modulate cortical responses with attentional focus than are NH listeners. Across individual listeners, performance on the selective attention task correlates both with ITD sensitivity and with the strength of attentional modulation of cortical responses. While, in NH listeners, attentional modulation of neural responses increases robustly over the course of a stream, this increase in attentional focus is weak in HI listeners. These findings show that hearing loss is associated with poor spatial acuity, which interferes with the ability to filter out sound sources based on their location, thereby contributing to communication difficulties that HI listeners face in common social settings.

Results

In the main selective auditory attention task, a visual cue at the start of each trial told listeners to focus attention on either the left, center, or right melody in an upcoming sound mixture. The mixture contained three competing, rhythmically regular melodies (the Distractor Stream, Leading Stream, and Lagging Stream; see Fig. 1*A*). The melodies started at different times and had different presentation rates, so that neural event-related potentials (ERPs) evoked by the onsets of notes in each stream (measured using EEG) could be isolated in time. The first melody to begin, the Distractor Stream, was a four-note melody that never was from the cued direction (always was to be ignored). The remaining two melodies comprised five notes (the Leading Stream, which began 490 ms after the Distractor) and four notes (the Lagging Stream, which began 200 ms after the Leading Stream). In each trial, one of the melodies was randomly chosen to have an ITD of zero, while the other two melodies were randomly assigned symmetrical, lateral locations in opposite hemifields. The ITDs of the two lateral streams were either small (±205 μs) or large (±799 μs), to vary the task difficulty (Fig. 1*B*). (Note that, to keep our language concise, we use the phrase "spatial configuration" to refer to specific choices of ITDs for the Leading, Lagging, and Distractor Streams, even though we did not present natural combinations of spatial cues.) Listeners were asked to report whether the melodic contour of the stream from the cued direction (the "target stream," which was either the Leading or the Lagging Stream) was rising, falling, or zig-zagging. EEG was measured throughout the task. (See *Materials and Methods* for more details about the stimuli and procedures.)

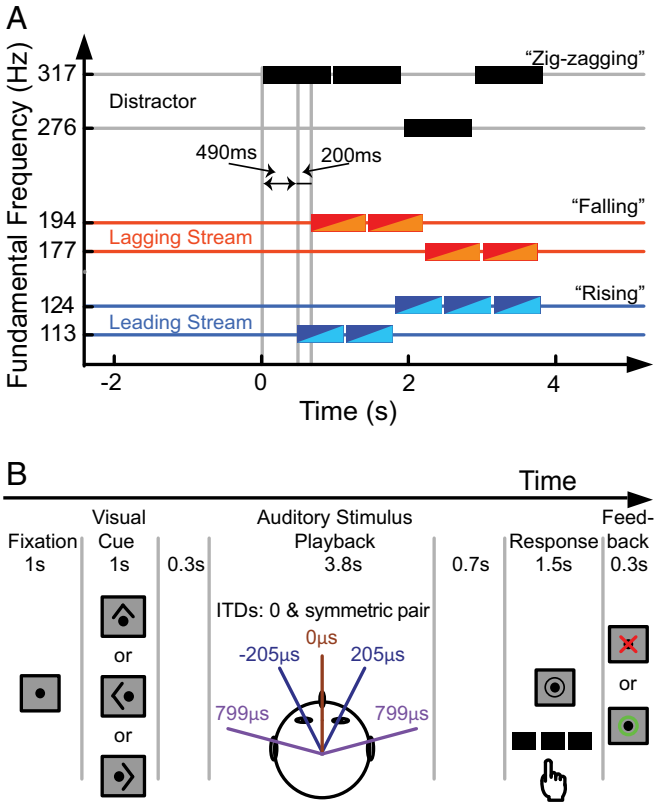

Fig. 1. Outline of the stimuli and experimental procedures. (*A*) Each trial presented three isochronous streams, each making up a melody of H and L notes. The onsets of the notes in each stream were staggered in time to allow EEG responses to be isolated in time. The Distractor Stream always began first, made up of complex tones with fundamental frequencies of 275 Hz and 317 Hz (composed of the first three harmonics at equal amplitude). The Leading Stream (five notes) began next, and consisted of complex tones of 113 Hz and 124 Hz. The Lagging Stream (four notes) started last and consisted of complex tones of 177 Hz and 194 Hz. Leading and Lagging Stream notes each contained the first 33 harmonics of their fundamental, at equal amplitude. Each melody either was rising, falling, or zig-zagging. (*B*) Each trial began with a fixation dot, followed by a visual cue, followed by the auditory stream mixture. The visual cue indicated the direction to attend in the upcoming auditory mixture (center, left, or right). The three auditory streams were then presented from three different lateral angles, simulated using only ITDs: center, and symmetrically on the left and right. Left and right streams either both had small or both had large ITDs, depending on the trial. Listeners were cued to identify the melodic contour of the target stream at the end of each trial, and were provided feedback after they responded.

HI Listeners Perform Poorly on a Spatial Selective Auditory Attention Task. All listeners are able to identify melodies presented in isolation (with no competing streams) with little trouble (*Materials and Methods*). However, in the mixture case, NH listeners are better than HI listeners at reporting the melodic contour of the target (Fig. 2*A*). Both groups do worse with the small ITD than with the large ITD. These observations are supported by the results of a planned multiway, repeated-measures ANOVA with factors of hearing status (NH vs. HI), ITD (small vs. large), and target stream (Leading vs. Lagging). Main effects of hearing status [$F(1, 37) = 13.57, P < 0.001$)] and ITD [$F(1,74) = 57.26, P < 0.001$] are significant, but not target stream [$F(1,37) = 0.0002, P = 0.988$]. The interaction between hearing status and target stream also reaches significance [$F(1, 37) = 6.028, P < 0.019$], but none of the other interactions do ($P > 0.05$ for all other interactions).

To explore the interaction between hearing status and target stream, we conducted follow-up repeated-measures ANOVAs separately for NH and HI listeners, with factors of ITD and target stream. For NH listeners, performance is significantly worse for the Leading Stream than for the Lagging Stream [$F(1, 23) = 8.157$, $P = 0.00894$] and significantly worse for the small ITD than the large ITD [$F(1, 46) = 43.04, P < 0.0001$], but the interaction is not significant [$F(1,46) = 0.0681, P = 0.795$]. For HI listeners, there is no significant effect of target stream [$F(1,14) = 1.716, P = 0.211$], but performance is significantly worse for the small ITD than the large ITD [$F(1,28) = 15.63, P = 0.00047$]; their interaction is not significant [$F(1,28) = 0.0955, P = 0.760$].

These analyses show that, overall, hearing impairment interferes with selective auditory attention. Performance for both NH and HI listeners is better when the ITD separation between streams is larger. NH listeners are better at attending to and identifying the melodic contour for the higher-pitched Lagging Stream compared with the Leading Stream, but this is not the case for HI listeners.

Performance Depends on the Spatial Focus of Attention. Listeners tended to do worse when the target stream was at midline compared with when it had an ITD to either the left or right. We did post hoc, paired two-sided *t* tests, separately for NH and HI listeners, to test whether this tendency was statistically significant.

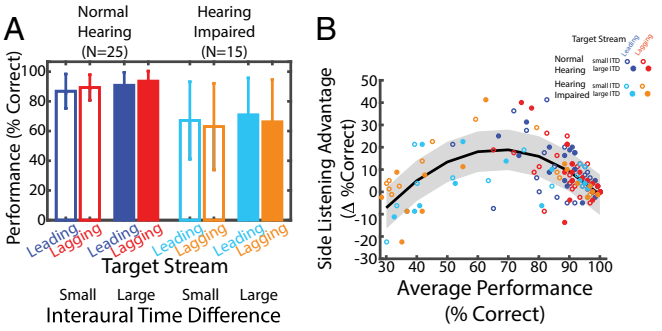

Fig. 2. HI listeners performed worse in a spatial selective auditory attention task than NH listeners. (*A*) Percent correct scores on the main attention task for both NH (left: red and blue bars) and HI (right: light-blue and orange bars), with across-subject SDs (note that, in this figure and throughout, intersubject differences are large, but consistent across conditions, leading to large error bars despite consistent effects of condition). Overall, NH listeners outperformed HI listeners. Within both groups, performance was better when sources were separated by a large ITD (filled bars) than a small ITD (open bars). For NH (but not HI) listeners, performance was better for the higher-pitched Lagging Stream (red bars) than the Leading Stream (blue bars). (*B*) The SLA, or difference in percent correct performance for the same spatial configuration when the target is on the side compared with when it is at midline, plotted as a function of the average performance for this spatial configuration across all target locations. Individual listener results for each main condition are shown by individual data points. Expected middle quartiles of the distribution, based on a second-order polynomial curve fit, are shown in gray. When near chance or near perfect performance, there is little effect of target location, but the SLA is generally positive for midlevel performance. Thus, individual subject ability in a given spatial configuration determines whether one sees evidence that performance for a target at midline, surrounded by distractors, is worse than performance when the target is to the side.

For each condition and subject, we compared performance when the target was at midline to average performance when the target was to either side of midline. For both NH and HI listeners, performance is significantly lower when the target is diotic (with competing streams to either side) compared with when it is from the side [NH: $t(99) = 7.413$, $P < 0.0001$); HI: $t(59) = 3.408$, $P = 0.0012$].

Looking at the results in more detail, floor and ceiling effects influence the dependence of performance on the direction of attention. Some subjects in both groups perform nearly perfectly in some conditions, while, in the hardest conditions, the worst HI listeners are near chance (33% correct), as if they are guessing, no matter what the target direction. For each listener in each of the four main conditions, we computed the difference between performance for a lateral target and performance for a central target, and looked at how this "side listening advantage" (SLA) varied with overall performance (Fig. 2*B*). We combined results for NH and HI listeners [(25 + 15) subjects × 4 conditions, for 160 independent data points)] and performed a second-order polynomial fit, predicting the SLA from the average level of performance across target locations (gray region in Fig. 2*B* shows expected range of 50% of the data). This analysis shows that the SLA is limited by floor effects near the guessing limit (33%) and by a ceiling as performance approaches 100%. Importantly, in the midrange of performance, the SLA is positive. In other words, when performance is not at floor or ceiling, listeners tend to perform better when listening for a lateral target compared with a central target.

HI Listeners Are Relatively Insensitive to Interaural Time Differences. Compared with HI listeners, NH listeners have lower (better) thresholds for detecting ITDs in both Leading Stream notes (18.4 ± 8.6 μs vs. 38.2 ± 15.8 μs) and Lagging Stream notes (28.5 ± 16.3 μs vs. 73.3 ± 62.0 μs), estimated using an adaptive procedure (see *Materials and Methods*). Similar to the behavioral results from the main selective attention task, these results show that there are significant differences in ITD sensitivity between the NH and HI groups.

Attentional Modulation of Cortical Responses Is Weak in HI Listeners. From the EEG data collected during the main task, we calculated an "attentional modulation index" (AMI) to quantify how the neural representation of Leading and Lagging Streams changes due to attentional focus. Specifically, the AMI is the normalized difference in the magnitude of ERPs evoked by notes in a stream when that stream is attended vs. when it is ignored (see *Materials and Methods*). In general, the AMI will be positive if responses are smaller when a stream is ignored compared with when it is attended; the larger the AMI, the greater the effect of attention on the neural representation.

For both NH and HI listeners, the first note in the Leading and Lagging streams shows no significant attentional modulation due to top-down focus (the distributions of AMI for the first notes in a stream are not significantly different from zero for either stream with either ITD; $P > 0.05$, Wilcoxon Signed Rank Test). Given this, we quantified the strength of attentional modulation, overall, for each listener in each condition as the average of the AMIs of notes 2 to 5 for the Leading Stream and of notes 2 to 4 for the Lagging Stream, in each of the ITD conditions.

Although intersubject differences in the AMI are large, the average AMI is generally smaller for HI listeners than for NH listeners (Fig. 3*A*). We performed a multiway, repeated-measures ANOVA with factors of hearing status (NH vs. HI), ITD (small vs. large), and target stream (Leading vs. Lagging) to analyze how AMI varies across groups and conditions. Main factors of hearing status [$F(1, 53) = 15.57$, $P = 0.000235$)] and ITD [$F(1, 74) = 4.136$, $P = 0.0456$)] are significant, as is the interaction between hearing status and ITD [$F(1, 74) = 25.12$, $P = 3.56 \times 10^{-6}$] and the three-way interaction between hearing status, ITD, and stream [$F(1,74) = 10.31$, $P = 0.00196$]. These results show that, as a group, NH listeners are better at modulating the cortical representation of a sound mixture when they focus spatial selective auditory attention than are HI listeners; however, performance of the two groups changes differently with condition.

To tease apart the various interactions, we undertook two separate repeated-measures ANOVAs on the AMIs for the two listener groups with main factors of ITD and target stream. For the NH listeners, ITD is significant [$F(1, 46) = 24.41$, $P = 1.0756 \times 10^{-5}$], stream trends toward significance [$F(1, 45) = 3.251$, $P = 0.0780$], and their interaction is significant [$F(1, 46) = 9.156$, $P = 0.0040$]. These results show that, for large ITD conditions, NH listeners have larger AMIs for the Lagging Stream than for the Leading Stream; in contrast, for the small ITD condition, NH listeners have larger AMIs for the Leading Stream than for the Lagging Stream. For the HI listeners, ITD is significant [$F(1, 28) = 6.888$, $P = 0.0139$]; however, stream is not significant [$F(1, 22) = 2.492$, $P = 0.1288$], and their interaction trends toward significance [$F(1, 28) = 3.387$, $P = 0.0763$]. These various effects reflect a trend for the AMI to be weakest for the large ITD, Lagging Stream condition, and similar across the remaining conditions.

Attentional Modulation Increases Across Time, but only for NH Listeners. The strength of attentional modulation tends to increase over the course of a stream, especially for NH listeners (Fig. 3*B*). Specifically, for NH listeners, the AMI increases with note position (Spearman's rho reached statistical significance for all four combinations of Leading and Lagging Streams and small and large ITDs; see text inserts within the panels of Fig. 3*B*); however, for HI listeners, the AMI correlation with note position only reaches significance for the Lagging Stream, small ITD condition. To visualize these effects, Fig. 3*B* plots the AMI of the first note and the last note in each of the four conditions, which were compared statistically using the Wilcoxon Ranked Sum test. For NH listeners, the AMI is always statistically

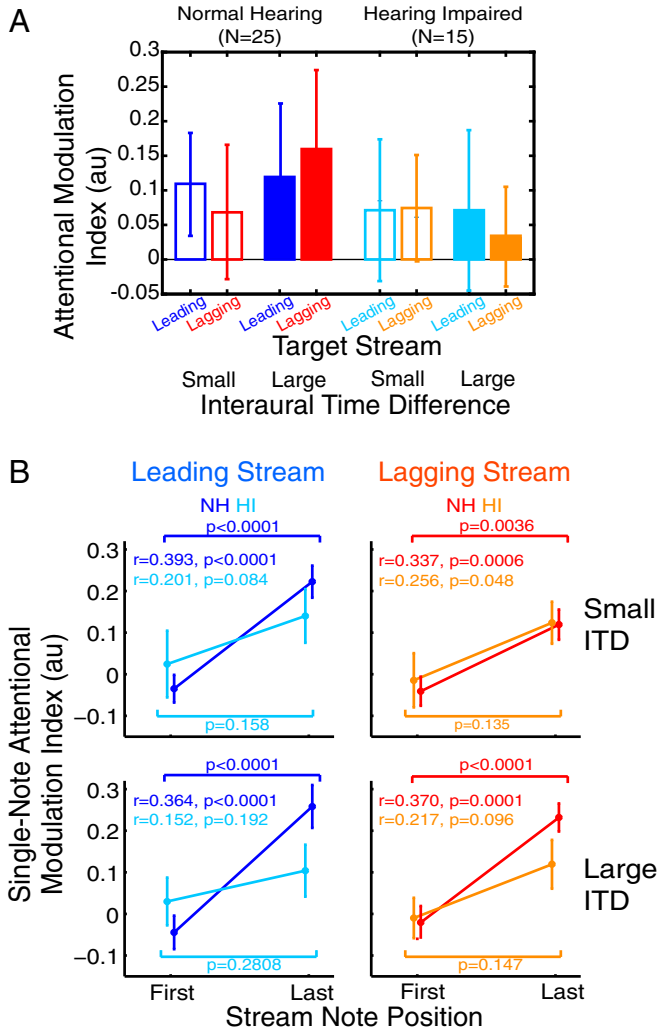

Fig. 3. NH listeners show weaker and less consistent modulation of neural responses with the focus of attention. (*A*) The AMI, or normalized change in the ERP magnitude due to shifts in attentional focus, is plotted using the same layout as in Fig. 2. Error bars show across-subject SDs. Overall, the AMI is greater in NH than HI listeners, as confirmed by a multiway ANOVA that finds hearing status to be a significant factor influencing AMI (see *Attentional Modulation of Cortical Responses is Weak in HI Listeners*). (*B*) NH, but not HI, listeners show a significant buildup of attention over the course of a trial. The AMI, or normalized change in the ERP magnitude due to shifts in attentional focus, is plotted for the first and last notes for both the Leading Stream (*Left*) and the Lagging Stream (*Right*), with SDs across subjects. Results for small ITDs (*Top*) and large ITDs (*Bottom*). (*Insets*) The Spearman's rho rank correlation between the AMI strength and the note position, from first to last. Whereas there is robust evidence that attentional modulation increases from note to note for NH listeners based both on differences in the AMI from first to last notes and from rank correlations (blue and red), there is little evidence of such effects in the HI listeners (light blue and orange).

greater for the final note than the initial note. However, for the HI listeners, the AMI is not significantly larger for the last note compared with the first note for any of the four conditions.

Attention Has No Effect on Neural Responses to the Initial Distractor Note. The visual cue telling listeners where to focus attention in an upcoming trial occurs 1.3 s before the first note of the Distractor Stream (the first sound in each trial). Previous work has shown preparatory neural activity once a listener knows where to attend but before the onset of an auditory target (23, 24). We asked whether the visual cue for where to attend sets up a "spatial filter" that modulates the ERP evoked by the first note of the Distractor (i.e., whether the ERP magnitude depends on the relative locations of the Distractor and the focus of attention). Initial analysis showed that responses were roughly symmetrical, left/right, so we calculated the ERP magnitude evoked by the first note of the Distractor for each of its possible locations (zero, small, or large ITD) and target directions (lateral or central; see Fig. 4*A*). Because the visual cue was "left," "center," or "right," but did not indicate whether the target ITD would be small or large, the listener state was identical for small and large ITDs at the onset of a diotic Distractor, before the Leading and Lagging Streams began (leftmost bars in panels of Fig. 4*A*); however, for lateral Distractors, the physical stimuli differed with ITD condition from the first note, leading to four additional configurations (depending on the Distractor ITD and whether the visual cue instructed listeners to attend to the opposite side or to the center; see two rightmost bar pairs in the panels of Fig. 4*A*).

ERPs to the first Distractor note vary significantly with Distractor location. Specifically, one-way ANOVAs, conducted separately for NH and HI listeners, reveal significant effects of Distractor ITD on the evoked ERP [NH: $F(2,98) = 18.89$, $P < 0.0001$; HI: $F(2,58) = 9.366$, $P = 0.0003$]. Post hoc Tukey tests showed that the ERPs were larger when the Distractor had a large ITD compared with when it was either zero or the small

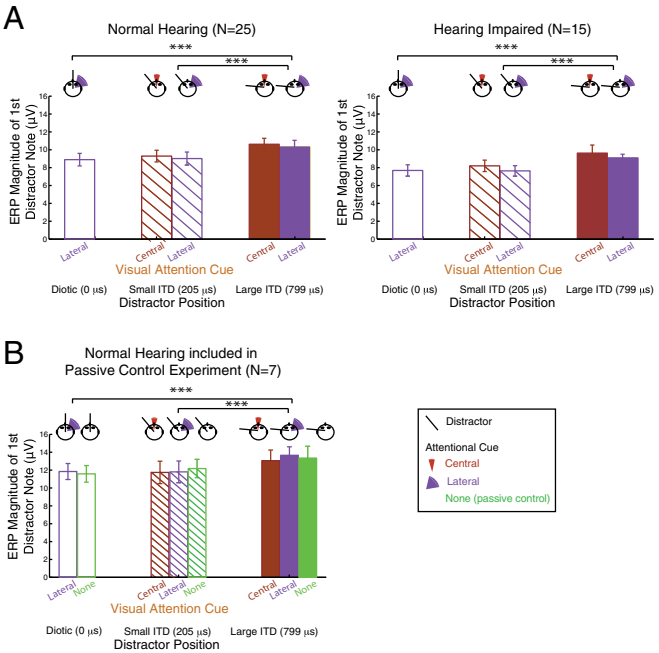

Fig. 4. The first note of the Distractor Stream is not modulated by spatial focus of attention for either NH or HI listeners. (*A*) The ERP magnitudes evoked by the first note of the Distractor in the main attention experiment are plotted for different spatial configurations, both for NH (*Left*) and HI (*Right*) subjects, with across-subject SDs. Because results were left–right symmetric, spatial configurations are combined for mirror-symmetric results. Each set of bars corresponds to a different Distractor Stream ITD (zero ITD, small ITD, large ITD, from left to right). Within each set of bars, results are broken down by the direction of attentional focus, denoted by the color wedge (maroon for central, purple for lateral). For both NH and HI listeners, the ERP for the large ITD Distractor has a greater magnitude than for the zero or small ITD. (*B*) ERP magnitudes for a subset of the NH listeners who completed a passive-listening control experiment, in a layout like that in *A*. The passive-listening data are shown in green, alongside the results from the main experiment for the same subset of subjects (included in *A*). The differences in ERP size seen in *A* are present in the passive-listening results, and thus can be attributed to differences in the Distractor ITD, rather than top-down attention.

ITD, but no other differences were significant (NH: z = 5.103 for large vs. zero ITD and z = 5.212 for large vs. small ITD, both with $P < 0.0001$; HI: z = 3.534, $P = 0.0008$ for large vs. zero ITD and z = 3.723, $P = 0.0006$ for large vs. small ITD). The ERPs for the zero ITD and small ITD Distractor did not differ significantly in either group ($P > 0.05$).

We reasoned that the large ITD Distractor may elicit a larger ERP than the other configurations either because of top-down attention effects or due to differences in the ERP magnitudes evoked by notes with different ITDs. To test this second possibility, we did a post hoc, passive control experiment with seven of our original NH listeners. In this test, the listeners watched a silent, subtitled movie. We presented individual notes with small and with large ITDs, but instructed listeners to ignore these sounds. The ERPs evoked in this passive case were statistically indistinguishable from those evoked by the same notes in the active attention task (one-way ANOVAs testing for effect of direction of attention yielded $P > 0.05$ for all Distractor ITDs; see Fig. 4*B*). These results thus show that there is no effect of attention on the first Distractor note, but that in quiet, notes with a large ITD elicit larger ERPs than the other notes.

HI Listeners Show Weak Suppression of Distractor Notes When Attending a Midline Target. In general, the ERPs evoked by later Distractor notes, which overlap temporally with the Leading and Lagging Streams, are much smaller than the ERPs evoked by the initial Distractor note. To quantify the strength of the response to the late Distractor notes, we averaged ERP magnitudes across notes 2 to 4 (Fig. 5). Given that our control experiment showed that the size of the ERP elicited by the initial note varies with location, we did a multiway ANOVA separately for each Distractor ITD (diotic, small ITD, and large ITD), with main factors of hearing status (NH and HI) and direction of attention (e.g., for a diotic Distractor, attention could be directed to a target with either a small or large ITD).

When the Distractor is at midline and attention is directed to a target to the side (leftmost bars in the panels in Fig. 5), there is no significant difference between NH and HI listeners [$F(1,38) = 1.613$, $P = 0.2118$]. For the HI listeners, the ERPs to the central Distractor are slightly larger when attention is directed to a target with a small ITD vs. a large ITD; however, there is neither a statistically significant effect of direction of attention [$F(1,38) = 0.5749$, $P = 0.4530$], nor a significant interaction between hearing status and direction of attention [$F(1,38) = 1.700$, $P = 0.2001$].

For both small ITD and large ITD Distractors, ERPs are significantly smaller for NH than for HI subjects [main effect of hearing status: $F(1,38) = 8.2211$, $P = 0.0067$ for small ITD; $F(1,38) = 6.7140$, $P = 0.0135$ for large ITD], showing that NH listeners are overall more effective at suppressing the Distractor. The main effect of direction of attention is not significant for either of the Distractor locations [small ITD: $F(1,38) = 2.728$, $P = 0.1068$; large ITD: $F(1,38) = 3.998$, $P = 0.0527$]; however, in both cases, there is a significant interaction between hearing status and direction of attention [$F(1,38) = 6.8939$, $P = 0.0126$ for small ITD; $F(1,38) = 6.9180$, $P = 0.0123$ for large ITD]. Post hoc tests show that, for NH listeners, the direction of attention has no impact on the Distractor ERP size for either of the Distractor locations [$F(1,24) = 0.1381$, $P = 0.7134$ for small ITD; $F(1,24) = 0.0011$, $P = 0.9734$ for large ITD]; however, for HI listeners, Distractor ERPs are larger when the target is at midline compared with when the target and Distractor are in opposite hemifields [$F(1,14) = 5.9095$, $P = 0.0291$ for small ITD; $F(1,14) = 7.9944$, $P = 0.0134$ for large ITD]. This difference shows that HI listeners are less effective at suppressing the Distractor when the target is from the center location compared with when the target is to the side.

Looking across conditions, all of the late Distractor notes elicit roughly the same magnitude response in NH listeners, even though the size of the ERP to the initial Distractor note varies with ITD. In other words, for NH listeners, the response elicited by late Distractor notes does not depend on the spatial configuration of the three-stream mixture; instead, the late Distractor notes are all suppressed equally well. In contrast, HI listeners do not suppress the late Distractor notes with small and large ITDs as well as NH listeners; importantly, this suppression is especially weak when they attend to a target at midline (in Fig. 5, *Bottom*, compare left and right for the two rightmost dyads of bars).

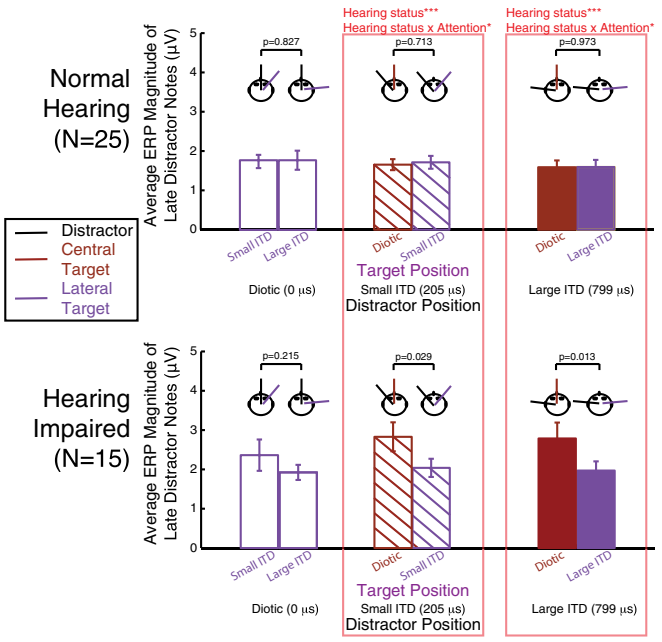

Fig. 5. NH listeners are equally good at suppressing later Distractor notes, independent of the spatial configuration, but HI listeners suppress Distractor notes only weakly when trying to focus attention on a midline target (with the Distractor Stream to one side). The average ERP magnitude evoked by later notes of the Distractor are plotted for NH (*Top*) and HI (*Bottom*) subjects, with across-subject SDs. Each set of bars corresponds to a different Distractor Stream ITD (zero ITD, small ITD, large ITD, from left to right). Within each set of bars, results are broken down by the direction of attentional focus, denoted by the bar color (maroon for central, purple for lateral) and labeled according to target ITD. HI listeners had larger ERPs, overall, for small ITD and large ITD Distractors (middle and right columns). In addition, in these configurations, HI listeners had significantly larger Distractor ERPs when the target was at midline compared with when the target was to one side. These results show that HI listeners tend to be worse at suppressing task-irrelevant Distractors than NH listeners, especially in the most challenging listening condition, when the target is at midline and surrounded by competing streams.

Individual Differences in Selective Attention Performance, Attentional Modulation, and ITD Sensitivity Are Consistent and Interrelated. When identifying a particular stream's melody, selective attention performances for the large and small ITD conditions are correlated across listeners in both listener groups (NH: $r = 0.9162$, $P < 0.0001$ for Leading and $r = 0.8836$, $P < 0.0001$ for Lagging; HI: $r = 0.9756$, $P < 0.0001$ for Leading and $r = 0.9910$, $P < 0.0001$ for Lagging). Moreover, the average ability to identify the Leading Stream melody is strongly correlated with the average ability to identify the Lagging Stream melody (NH: $r = 0.8949$, $P < 0.0001$; HI: $r = 0.8888$, $P < 0.0001$). This analysis shows that there are very consistent individual differences in selective attention performance across conditions, even though performance varies across conditions, as noted in *Performance Depends on the Spatial Focus of Attention* when considering group-level effects.

Similarly, individual differences in the average AMIs are consistent across conditions for the listeners in each group. For both streams, the AMIs for large and small ITD conditions are

correlated (NH subjects: $r = 0.7164$, $P < 0.0001$ for the Leading Stream and $r = 0.5904$, $P = 0.0019$ for the Lagging Stream; HI subjects: $r = 0.8787$, $P < 0.0001$ for the Leading Stream and $r = 0.6129$, $P = 0.0151$ for the Lagging Stream). After collapsing across ITD conditions, average AMIs for the Leading and Lagging Streams are correlated (NH: $r = 0.5219$, $P = 0.0075$; HI: $r = 0.6275$, $P = 0.0123$). This analysis shows that, just as with performance, there are very consistent intersubject differences in the effects of attention on neural responses, even though the AMI changes with condition.

Within both the NH and HI groups, behavioral performance correlates with the AMI (Fig. 6). In particular, the correlation between performance and AMI is significant for both groups for the Leading stream, small ITD condition (NH: $r = 0.6604$, $P = 0.0003$; HI: 0.7639, $P = 0.0009$), the Leading stream, large ITD condition (NH: $r = 0.6439$, $P = 0.0005$; HI: $r = 0.7805$, $P = 0.0006$), and the Lagging stream, small ITD condition (NH: $r = 0.5637$, $P = 0.0033$; HI: $r = 0.6153$, $P = 0.0146$); also, for both groups, the correlation between performance and AMI trends toward, but does not reach significance for, the Lagging Stream, large ITD condition (NH: $r = 0.3933$, $P = 0.0517$; HI: $r = 0.5128$, $P = 0.0506$). In all four conditions, NH listeners with low AMIs tended to perform better than corresponding HI listeners, leading to a stronger dependence of performance on AMI for HI listeners. Indeed, NH listeners whose AMI was near zero generally performed well above chance levels, while the HI listeners who did not show any attentional modulation of cortical responses tended to perform near chance (33% correct) in each of the conditions. This analysis shows that there is a strong relationship between performance and attentional modulation of neural responses for both NH and HI listeners. Notably, however, even the worst NH listeners are still able to perform relatively well on the selective attention task, even though they may show almost no attentional modulation.

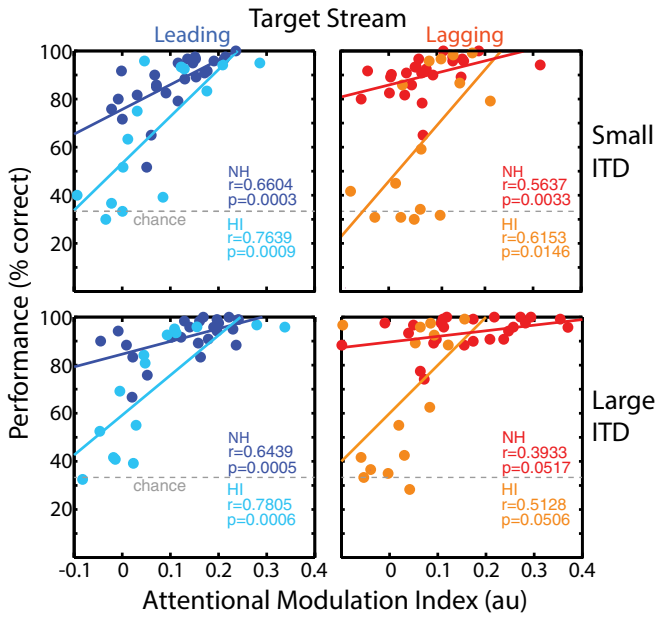

Fig. 6. For both NH and HI listeners, an individual's performance on the selective attention task correlates with the strength of their attentional modulation, showing that listeners who perform well on the attention task show stronger modulation of neural responses based on attentional focus. Each of the four panels shows a scatterplot of individual subjects' performance as a function of their AMI for a given condition. Plotted results when attending to (*Left*) the Leading Stream and (*Right*) the Lagging Stream. Results for (*Top*) the small ITD configurations and (*Bottom*) the large ITD configurations. Within each panel, results are shown for both NH (primary colors) and HI (secondary colors) listeners. Regression lines of the appropriate color show the relationship between AMI and performance for the corresponding listener group in that condition. It is noteworthy that the worst HI listeners are near chance levels on the task, and show no significant attentional modulation; however, even though the worst NH listeners may show no significant attentional modulation, they nonetheless perform well above chance on the attention task.

In addition, in the HI listener group, there was a trend for the pure-tone average threshold to be negatively correlated with AMI; that is, the HI listeners with the worst hearing tended to have the weakest modulation of neural responses based on attentional focus. However, for our cohort of HI listeners, this trend did not reach statistical significance ($P = 0.4781$, $r = 0.0714$; see Fig. S2*A*).

Individual differences in ITD sensitivity are also consistent across subjects. Although ITD thresholds are generally higher for the higher-pitched Lagging Stream than the lower-pitched Leading Stream, within both groups, ITD thresholds for the Leading and the Lagging notes are strongly correlated (NH: $r = 0.7729$, $P < 0.0001$; HI: $r = 0.9563$, $P < 0.0001$). This result shows that individual differences in ITD sensitivity are consistent across the two streams. Within the HI listener group, there is a significant correlation between the average ITD threshold and pure-tone average hearing threshold ($P = 0.8079$, $r = 0.0003$; see Fig. S2*B*). Thus, for HI listeners, the degree of hearing loss predicts sensitivity to ITDs.

The sound mixtures presented in the spatial selective attention task always contained a mixture of Leading and Lagging sources. Individual differences in ITD sensitivity to Leading and Lagging sources are correlated. Therefore, to examine whether ITD sensitivity is related to overall performance on the selective attention task, we compared each subjects' average ITD threshold to their grand average performance on the task (Fig. 7). For both groups, there is a significant relationship between performance and ITD sensitivity (NH: r = –0.5960, $P = 0.0017$; HI: r = –0.5837, $P = 0.0223$).

From these examinations of individual differences, we see that performance on a spatial selective attention task, the ability to steer attention to the auditory target, and ITD sensitivity are all intimately linked. Moreover, for HI listeners, the degree of hearing loss is related to ITD sensitivity.

Discussion

HI Listeners Have Poor Spatial Acuity, Poor Spatial Selective Auditory Attention, and Weak Attentional Modulation. NH listeners, as a group, demonstrate better spatial acuity, have better selective attention performance, and show larger changes in neural responses based on attentional focus than HI listeners. In addition to these group differences, on an individual level, differences in performance on the attention task are correlated both with neural measures of attentional modulation and with ITD sensitivity.

These results are similar to past results in NH listeners showing that attentional modulation of cortical responses predicts individual performance on selective auditory attention tasks (18, 22, 25). The relationship between attentional modulation and spatial acuity found in the current study suggests that spatial acuity determines how "selective" attention to a stream from a particular location can be, which then impacts how strongly other streams from other locations will be suppressed. In our study, we also find that hearing loss correlates with ITD thresholds; that is, listeners with the greatest hearing loss tend to have the worst ITD sensitivity. If a listener has hearing loss that produces poor spatial acuity, they will suppress interfering streams only weakly, making it harder for them to analyze the target stream. This helps explain some of the communication difficulties that listeners with hearing loss experience in noisy social settings.

Effects of Hearing Loss on Selective Attention Performance Depend on the Experimental Details. Hearing loss is known to degrade performance on spatial selective auditory attention tasks using competing speech streams (2–4). We find similar results using a

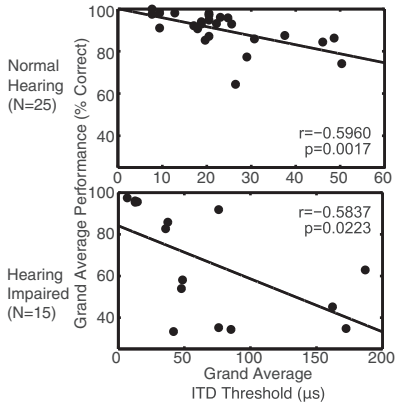

Fig. 7. For both NH and HI listeners, an individual's overall performance level on the selective attention task is negatively correlated with their average ITD threshold. Scatterplots for NH listeners (*Top*) and HI listeners (*Bottom*), showing overall percent correct as a function of ITD threshold. Best-fit regression lines are shown as lines in each panel. Note that the *x* axes cover a much larger range for the HI listeners than for the NH listeners. For both groups, overall performance is correlated significantly with performance.

task in which listeners identified simple melodies in the presence of competing melodies from different lateral positions. Previous studies have also demonstrated that hearing loss is associated with poorer binaural acuity (1), similar to what we report here. However, past efforts to document relationships among hearing status, binaural acuity, and spatial selective auditory attention in speech mixtures have met with mixed results: While some have found correlations among these measures (5), the correspondence is not always clear (6).

The current results help to resolve some of these apparent discrepancies. Our EEG results show that HI listeners suppress a lateral Distractor less than do NH listeners when the target stream is at midline. In contrast, our NH listeners are good at suppressing the Distractor no matter what the lateral position of the target or its position relative to the Distractor. Behaviorally, we see that both NI and HI listeners are generally poorer when a target is at midline compared with when it is to one side; that is, the most challenging spatial configuration, behaviorally, is the condition where HI listeners show the greatest deficits in neural suppression. However, the size of the behavioral SLA depends on the general level of performance, with little effect near floor or near ceiling.

Together, these results are consistent with the idea that hearing loss, which produces poor spatial acuity, is associated with a "broader" attentional spatial filter. In turn, a broader spatial filter produces weak segregation and suppression in some, but not all, listening conditions. For instance, if competing sound sources are separated by more than a listener's "attentional beam width," then spatial acuity will be a poor predictor of performance, which will be good in general. Similarly, if the competing sources are so close together that even the best listeners have trouble resolving them, then spatial acuity will also fail to predict performance. Given this logic, whether or not one finds a clear relationship between spatial acuity and the ability to understand speech amid spatially separated competitors will depend on the exact spatial configurations and population of listeners tested.

In the current study, the adjacent streams had ITDs that differed by either 205 μs (small ITD) or 699 μs (large ITD). ITD thresholds in our NH group were all under 50 μs, but ranged from under 50 μs to near 200 μs in our HI group (Fig. 6). As these thresholds are the smallest ITDs that can be reliably detected, they are undoubtedly smaller than the "width" of spatial attentional focus. Given that ITD discrimination thresholds for many of our HI listeners were of the same magnitude as the ITDs separating the competing sources, it is no surprise that many did quite poorly on the spatial selective attention task, with some performing at chance levels. While none of our NH listeners were at chance performance, the small ITD separation was still small enough to produce some perceptual interference for many listeners. Thus, because the small ITD separation tested here was "on the edge" of what NH listeners can reliably use when focusing spatial attention, our results revealed strong correlations between spatial acuity and performance.

Attentional Focus Increases over Time for NH but Not HI Listeners. A spatial cue directing a subject's attention toward the location of an upcoming visual target causes anticipatory modulation of responses in visual cortex, even before the visual target occurs (26–30). Evidence of preparatory activity in anticipation of an upcoming target has been shown in previous auditory studies, as well (23, 24). Such results suggest that responses evoked by a subsequent sound will be relatively strong when the incoming sound matches the expected attributes of the target, and relatively weak when there is a mismatch. In the current study, a visual cue for where to attend was provided at the start of each trial. However, we find no evidence of spatially tuned suppression in the responses evoked by the leading note of the Distractor (Fig. 4). Instead, the response to the first Distractor note varied only with absolute spatial location, with lateral notes evoking stronger responses, consistent with previous reports (31). This suggests that any top-down filtering is overridden by the Distractor onset, which is inherently salient. The idea that sudden onsets are inherently salient enough to override any top-down effects of attention may also explain why neither NH nor HI listeners show evidence of any significant attentional modulation of the first notes of either the Leading or Lagging Streams.

For NH listeners, we find that the strength of attentional modulation increases over time, reducing interference from competing streams with each subsequent note. This increase in suppression may be the neural correlate of the observation that the ability to analyze an attended speech stream improves over time in the presence of spatially separated distractors in NH listeners (32, 33). Such "buildup of attention" likely reflects the gradual formation and refinement of auditory objects through time, which provides the substrate for selective attention (12, 34). These results suggest that, at the start of a trial, listeners roughly steer attention to the expected location of an upcoming target, and then hone their attentional focus once they begin to hear the target.

Importantly, for our HI listeners, selective attention strength increases weakly, if at all, over time. This failure to build up suppression suggests that HI listeners may not have adequate spatial resolution to separate the target content from interfering sound content: Their focus of attention is so coarse, they cannot isolate enough of the target to refine attention further.

Caveats: Experience, Audibility, and Stimulus Salience May Impact Performance for HI Listeners. We elected to test our subjects over headphones, without using hearing aids, and did not provide any frequency-specific compensation for hearing loss. We set the overall sound intensity of our three streams to be equal, and we set overall presentation levels to ensure the streams in our main task were clearly audible (*Materials and Methods*). We cannot be sure exactly how these choices impacted results, but it is important to consider how they may have.

We did not explicitly ask our HI listeners whether or not they normally wear their hearing aids; however, informally, based on our interactions with them, we estimate that less than half of them wear their hearing aids regularly. We also did not note any clear correspondence between task performance and regular hearing aid use. However, for listeners who normally wear aids, this experience may affect how they process spatial cues. Some listeners may wear an aid in only one ear, even though they have roughly symmetric losses, leading to an asymmetry in the sound

levels that they hear every day compared with what they heard in the laboratory. For listeners who typically wear older or less-expensive bilateral hearing aids, the two aids generally process the signals reaching the left and right ears independently, leading to inconsistent and fluctuating spatial cues; listeners who regularly wear such aids may have learned not to "trust" interaural difference cues. Alternatively, listeners wearing newer bilateral aids that tether processing across the ears to preserve interaural cues may be used to hearing very reliable spatial cues that are audible across a large range of frequencies. Thus, depending on the aids a listener uses every day, the weight they have learned to give to different spatial cues may vary. The question of how hearing aid usage—with different types of hearing aids—affects spatial processing is important and deserving of further attention. Still, our study shows that simple ITD thresholds are larger in HI listeners with greater hearing loss, and that poorer ITD thresholds are associated with poorer performance in a spatial selective attention task.

There are studies showing that interaural sensitivity is reduced at low sensation levels (SLs; sound levels re detection threshold) (35–37), which may have contributed to the weaker ITD sensitivity we observed in our HI listeners. Thus, providing amplification to increase stimulus SLs might improve ITD sensitivity for our HI listeners. It would be interesting, in future work, to determine whether there is such an improvement with increased SL, and, if there is, whether it translates into improved selective auditory attention and attentional modulation in HI listeners. Certainly, there are examples in the literature demonstrating that increased audibility can improve the intelligibility of speech presented against spatially separated interferers (7); however, because speech intelligibility itself will improve as audibility improves, there is no way to know if this is a spatial effect. A nonspeech paradigm like ours may help tease apart whether such improvements are due to enhanced spatial attention or simply better transmission of speech information.

There is another potential effect of audibility in the present study that deserves some consideration. Our HI listeners all had sloping hearing losses. The Leading, Lagging, and Distractor Streams differed in spectral content: The Distractor Stream content ranged from 276 Hz up to 951 Hz, the Leading Stream content ranged from 113 Hz up to 4,090 Hz, and the Lagging Stream content ranged from 177 Hz up to 6,402 Hz. As a result, the different streams were differentially affected by hearing loss. Indeed, each stream ended up being at a different SL for our HI listeners.

Differences in spectral content of the Leading and Lagging Streams, and how spectral content interacts with sloping hearing loss, may explain a subtle difference between performance for NH and HI listeners. We find that NH listeners perform better for the high-pitched, Lagging Stream than for the lower-pitched Leading Stream. This suggests that, for the NH listeners, the Lagging Stream is inherently more salient and easier to attend than the Leading Stream. In contrast, our HI listeners do not show this effect. However, because of their hearing loss, poor audibility of the high-frequency components in the Lagging Stream may reduce its inherent salience for our HI listeners, negating the effect seen in NH listeners.

The Distractor Stream has less high-frequency content than either the Leading Stream or the Lagging Stream, with the highest component below 1,000 Hz. As a result, compared with NH listeners, our HI listeners may find the Distractor Stream relatively more salient and more distracting, which may contribute to their overall poorer performance on the spatial selective attention task. For instance, the worst NH listeners showed almost no attentional modulation, yet were still able to perform well above chance levels, while our worst HI listeners were at chance. It may be that, in cases where "online" attentional modulation fails, listeners make judgments based on the raw, unmodulated sound mixture represented in the brain. In cases where one sound source in that raw mixture is dominant, but not the "target" source, the inherent dominance of the nontarget source may make it impossible for a listener to understand the target (e.g., for an HI listener for whom the Distractor Stream is relatively salient). However, if the competing sources in the raw sound mixture are more equal in their salience, a listener may still be able to pull out information from the target, especially if the target sound has audible components that are not overlapping spectrally with competing sounds in the mixture (e.g., for a NH listener, where the high-frequency components of the Lagging Stream are audible and not masked by any of the competing sounds).

However, differences in salience of the competing streams cannot explain all of the effects of hearing loss we find. In particular, we manipulated laterality of the competing streams using only ITDs; therefore, any differences in performance or in attentional modulation with the spatial configuration of the streams cannot be explained by salience, since the spectral content of the mixture reaching the ears is the same for all configurations. We find that, like NH listeners, HI listeners perform better when the stimuli are separated with large ITDs than with small ITDs (Fig. 2*A*) and that listeners who are not at floor are better when the target is to the side than when it is at midline (Fig. 2*B*). Importantly, as a group, HI listeners are as good as NH listeners at suppressing the Distractor when the target is to the side, even for the relatively low SL streams they are hearing; however, when HI listeners are trying to attend to a target stream at midline, they suppress a Distractor with identical spectral content less completely than do NH listeners. These results show that the spatial configuration of the streams, not just their spectral content or loudness, influences both performance and neural suppression of competing streams in our HI listeners. These differences between NH and HI listeners thus cannot be due solely to differences in inherent salience of the stimuli, but instead must reflect, at least in part, differences in the efficacy of spatial processing in the two groups.

We find a close correspondence between broadband ITD thresholds, the strength of attentional modulation of neural signals, and performance. These relationships, coupled with the fact that HI listeners show particular deficits in performance for some spatial configurations of the sources, support the view that HI listeners suffer from reduced binaural acuity, which impairs selective attention performance. Still, further experiments are necessary to tease apart how differences in overall salience of the streams comprising a sound mixture contribute to differences between NH and HI listeners.

Technologies That Rely on EEG Signatures of Attention May Face Challenges in HI Listeners. There has been a recent surge of research showing that attentional focus can be decoded from neural signals (18–21), but the vast majority of this research has been conducted on listeners with normal hearing. Our results complement those of one recent study showing that hearing loss negatively impacts cortical tracking of an attended speech stream in a two-talker mixture (38). Like this other study, we find that the EEG responses in HI listeners are less influenced by attentional focus than those in NH listeners. Moreover, the HI listeners who perform least well in the spatial selective attention task show little or no attentional modulation of their neural responses.

This observation has important implications for future technology development. Specifically, there has been increasing interest in developing hearing aids that incorporate EEG signals as an input to guide signal processing in complex scenes (19, 39–41). Unfortunately, this approach relies upon hearing aid users effectively modulating their neural responses based on attentional focus in order for decoding algorithms to read out cognitive intent. The current study suggests that the very listeners who are the most in need of assistance from such listening devices show almost no neural signatures of attention, making the problem of developing a useful "cognitively guided hearing aid" even more challenging.

Materials and Methods

Participants. NH control human subjects had thresholds of ≤20 dB hearing level (HL) at frequencies from 250 Hz to 8,000 Hz for both ears. Initially, 27 NH listeners were recruited; however, two of these failed the screening test (see *Auditory Selective Attention Procedures*); reported NH results are from the remaining 25 subjects (13 males, 12 females, aged 20 y to 52 y). HI subjects all had bilateral symmetrical sensorineural hearing loss, with pure tone thresholds of ≥25 dB HL at one or more frequencies between 250 Hz and 8,000 Hz and differences in thresholds across the two ears of ≤20 dB at every frequency. Fifteen HI listeners participated in the study (eight males, seven females, aged 20 y to 59 y). (See Fig. S1 for audiograms of the NH and HI listeners.)

Subjects gave informed consent as approved by the Boston University Charles River Campus Institutional Review Board, which also approved all of the experiments in the study. All subjects were compensated at an hourly rate and were paid an additional $0.02 bonus for each correct response in the main task to encourage them to remain attentive.

NH and HI groups did not differ significantly in age (two-sided Wilcoxon Rank Sum test; rank sum = 329, $P = 0.5651$) or in general attention ability, assessed by performance on two visual tasks from the Test of Everyday Attention (Pearson). Specifically, there was no significant difference between NH and HI groups on their scores in either the Map Searching Game p[$t(38) = 0.5405$, $P = 0.5920$ for 1-min score; $t(38) = 1.3111$, $P = 0.1975$ for 2-min score] or the Visual Elevator Game [$t(38) = -0.1562$, $P = 0.8767$ for correct number counted] using two-tailed unpaired t tests.

Equipment. Subjects sat in a sound-treated booth with a PC keyboard and monitor. The PC controlled the experiment using Psychtoolbox 3 (42) and MATLAB (Mathworks). Control code generated triggers marking key event times. Auditory stimuli were generated in MATLAB using a sampling rate of 48.828 kHz except for the ITD threshold test, which used 97.656 kHz. Signals were presented through a TDT System 3 unit (Tucker-Davis Technologies) and ER-1 insert headphones (Etymotic).

A BioSemi Active Two System recorded EEG signals with a 4.096-kHz sampling rate in 32 active scalp electrodes in the standard 10/20 configuration. Two additional electrodes were placed on the mastoids. EEG recordings were rereferenced to the mean of the two mastoid electrodes. Synchronized triggers from the TDT System were recorded alongside EEG data, which was stored on the controlling PC.

Auditory Selective Attention Procedures. In the auditory selective attention experiment, each trial consisted of three simultaneous, isochronous sequences of complex tones ("notes"; see Fig. 1*A*). Sequences differed in repetition rate, so that EEG responses elicited by onsets of the notes were resolvable in time. Two of the three streams could be the target (the Leading Stream and the Lagging Stream), while the other stream was always to be ignored (the Distractor). The Distractor, which started first, was made up of four notes, each of duration 919 ms, separated by an interstimulus (onset to onset) interval (ISI) of 959 ms. The Leading Stream started 490 ms after the Distractor and consisted of five notes, each of duration of 624 ms, separated by an ISI of 664 ms. The Lagging Stream started 200 ms after the Leading Stream and contained four notes of duration of 728 ms, separated by an ISI of 768 ms.

All notes were gated on and off with cosine-squared ramps (onsets: 10-ms duration; offsets: 100-ms duration) to reduce spectral splatter. Each note in the Leading and Lagging Streams consisted of the first 33 harmonics of some fundamental frequency, with equal amplitude, added in sine phase. Distractor notes were made up of the first three, equal-amplitude harmonics of their fundamentals added in sine phase.

For NH subjects, notes were played at 70 dB sound pressure level (SPL). Level was individually adjusted for HI listeners: Starting from 70 dB SPL, an example stream was played at increasing levels (steps of 5 dB) until it was comfortably loud. Five of the 15 HI subjects settled on 75 dB SPL, while the remaining 10 selected 70 dB SPL. We measured absolute threshold for each listener for each stream, and verified it was at least 10 dB SL. For the NH subjects, stream levels ranged from 50.00 dB SL to 61.00 dB SL (mean 57.00 dB SL) for the Leading Stream, 48.33 dB SL to 60.00 dB SL (mean 56.60 dB SL) for the Lagging Stream, and 48.33 dB SL to 51.00 dB SL (mean 50.65 dB SL) for the Distractor. For the HI subjects, the levels were 10.67 dB SL to 53.33 dB SL (mean 33.49 dB SL) for the Leading Stream, 10.67 dB SL to 52.33 dB SL (mean 32.82 dB SL) for the Lagging Stream, and 12.00 db SL to 51.00 dB SL (mean 36.02 dB SL) for the Distractor.

Each stream was a simple melody comprising low (L) and high (H) notes. L and H fundamental frequencies differed for each stream (Leading: 113.00 Hz 123.94 Hz; Lagging: 176.87 Hz and 194.00 Hz; Distractor: 275.96 Hz and 317.00 Hz), with pitch shifts of 1.6 semitones for Leading and Lagging Streams, and 2.4 semitones for the Distractor. On each trial, each stream was randomly and independently assigned a melody contour that was ascending, descending, or zigzagging (1/3 likelihood each). Ascending streams started with an L note, while descending streams started with an H note. Zigzagging streams could start with either note. For ascending and descending sequences, the melody changed value at some random point, and all subsequent notes repeated that value. For zigzagging melodies, at least two notes from the end of the melody, the note value changed; to ensure that listeners had to maintain attention on throughout the sequence, zigzagging melodies changed back at the final note.

The streams were simulated from three different locations, using ITDs. One of the streams was always diotic, while the other two were displaced symmetrically using either small (±205 μs) or large (±799 μs) ITDs, randomly, for each trial. All combinations of small or large ITDs, Leading Stream or the Lagging Stream as the target, and different spatial configurations of the three streams were tested.

Each trial started with a 1-s fixation dot, followed by a 1-s visual cue (Fig. 1*B*): an arrowhead around the fixation dot, pointing either left, right, or upward to indicate target location (left, right, or center, respectively). After a 0.3-s quiet period, the 3.8-s auditory stimulus was presented, followed by a 0.7-s silence. A circle then indicated the response period (1.5 s). Listeners were instructed to maintain eye gaze on the fixation dot, and to identify the target melody contour with a button press. The final button within the response time was the registered answer (no response was scored as incorrect). A 0.3-s symbol provided feedback (green circle or red X, respectively). The next trial began after a brief random pause (0 s to 0.1 s).

Before formal testing, each subject underwent training and screening to ensure that they could properly identify melodies in quiet. Initial training consisted of two 12-trial blocks of a single diotic stream (Leading Stream examples in the first block and Lagging Stream examples in the second block). Each subject then performed additional blocks until they achieved at least 8 of 12 correct for seven consecutive blocks. Two NH subjects failed to achieve this criterion and were excluded.

After preliminary training and screening, each subject completed 10 blocks of 48 trials of the main task (total of 480 trials). Each block consisted of two trials for each combination of ITD separation (small or large), target stream (Leading or Lagging), and spatial configuration (six different configurations), presented in random order.

ITD Threshold Procedures. ITD thresholds were measured adaptively for each subject with a three-down-one-up tracking procedure, using a step size of 10.24 μs. Thresholds were measured in two blocks, one using a 113-Hz Leading Stream note, the other a 194-Hz Lagging Stream note. Each trial started with a 0.6-s fixation dot followed by a 1-s pause. Two notes, separated by 0.4 s, were then played. The first was either left or right, with an ITD determined by the adaptive track, and the second was diotic. Subjects reported whether the first note was left or right of the second by pressing 1 or 2 during the response period (denoted by a circle around the fixation dot). After the response, feedback was provided by a 0.3-s-long symbol over the fixation dot (green circle or red X). ITD thresholds were estimated as the average of ITDs in the last six reversals of a total of nine. To qualify for the main experiment, thresholds for both targets had to be smaller than 205 μs (small ITD). No subject was rejected for failing to reach this criterion.

Passive Event-Related Potentials Procedures. We tested seven of the subjects from the NH group (three males, four females, aged 23 y to 43 y) in a passive EEG control experiment to explore whether ERP magnitudes depend on note ITD. Listeners heard 600 trials, each consisting of a single Distractor note (70 dB SPL; half L and half H). Notes were presented with ITDs of −799, −205, 799, and 205 μs (50 trials each for L and H) and 0 μs (100 trials each), for a total of 600 trials, in random order. The ISI was randomly selected from a uniform distribution from 3.9 s and 4.0 s. During the EEG recording, subjects watched a silent movie with subtitles and were asked to stay still and ignore the notes.

EEG Analysis. Raw scalp-recorded EEG signals were first band-pass filtered from 1.8 Hz to 50 Hz using a 6,000-point finite-impulse response band-pass filter (least-squares brick-wall filter; *firls.m* in MATLAB). We calculated ERPs from responses on channel Cz (channel 32 in the 10/20 system). Epochs were extracted from each raw trace, then band-pass filtered from 2 Hz to 25 Hz using the *eegfiltfft.m* function in EEGLab toolbox (43). Any trial epoch with a peak magnitude greater than 90 μV was rejected to remove artifacts (roughly 3 to 6% of trials were rejected).

We used a bootstrap procedure to estimate evoked EEG responses. We took the average of 100 randomly chosen trial epochs, with replacement, chosen from the distribution of artifact-free responses for a given subject and condition. The final estimate was the mean of 200 repeats of this procedure. From these estimates, the N1 magnitude was estimated as the local minimum 90 ms to 220 ms after a note onset, and the P1 magnitude was estimated as

the local maximum 30 ms to 120 ms after the onset. The difference in these magnitudes was used to quantify the early neural response for each condition and subject.

For each trial in the auditory selective attention experiment, epochs were defined from 3 s before to 6 s after the auditory stimulus onset was extracted. For the ITD threshold experiment, each trial epoch was defined as 0.5 s before to 1 s after the auditory stimulus onset. For each subject, we calculated the ERP magnitudes separately for large ITD, a small ITD, and 0 ITD trials.

Quantification of Attention in the Auditory Selective Attention Experiment. For the same stimulus, the P1–N1 magnitudes evoked by a note in a stream tended to be larger when listeners were attending to its stream than when they were attending to a different stream. We calculated the AMI for each note in each stream, for each listener and condition, as

$$\mathrm{AMI} = \frac{1}{\mathrm{First}}\left(N^{\mathrm{Attended}} - N^{\mathrm{Ignored}}\right), \quad [1]$$

where First denotes the average P1–N1 magnitude elicited by the first note of the Distractor, averaged across all trials for a particular subject, and N^{Focus} denotes the P1–N1 magnitude elicited by a particular note in a particular attention condition, which is either Attended or Ignored. Thus, the AMI is a normalized difference of the neural response to a stream when it is attended and ignored. The normalization, dividing by the average P1–N1 magnitude for the first note of the Distractor, reduces intersubject variability, which can be substantial. Because the first note of the Distractor is the first sound on each trial and is heard alone, it tends to evoke the largest P1–N1 response, and thus is a proxy for the largest P1–N1 magnitude a given listener produces.

After the first note, Distractor notes overlapped with other streams and tended to produce smaller ERPs. For these reasons, we separately analyzed responses to the first note and the later notes of the Distractor. We divided responses according to the Distractor's spatial location (diotic, small ITD, or large ITD) and the target stream location. We compared the P1–N1 peak-to-peak ERP magnitudes across conditions to analyze effects of attention on the Distractor.

Statistical Tests. Correlation coefficients were evaluated by Pearson's r or Spearman's rho, as noted. Most factors were analyzed using single-factor ANOVAs and multiway ANOVAs based on mixed-effects models (44), implemented in R (Foundation for Statistical Computing). Subject-related factors were treated as random effects. All other factors and interactions were treated as fixed-effect terms (although some factors were nested, precluding inclusion of some interaction terms). Post hoc Tukey tests explored relationships between subgroups of factors using Holm's method for *P* value adjustment. Some hypotheses were tested with two-tailed unpaired *t* tests or Wilcoxon Ranked Sum tests, implemented in R. All data sets subjected to parametric statistical tests were checked for normality using the Kolmogorov–Smirnov test.

ACKNOWLEDGMENTS. This work was supported by the Oticon Foundation and the National Institute for Deafness and Communication Disorders (Grants DC01598 and DC015760).

1. Colburn HS (1982) Binaural interaction and localization with various hearing impairments. *Scand Audiol Suppl* 15:27–45.
2. Mackersie CL, Prida TL, Stiles D (2001) The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss. *J Speech Lang Hear Res* 44:19–28.
3. Marrone N, Mason CR, Kidd G, Jr (2008) The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *J Acoust Soc Am* 124:3064–3075.
4. Best V, Marrone N, Mason CR, Kidd G, Jr, Shinn-Cunningham BG (2009) Effects of sensorineural hearing loss on visually guided attention in a multitalker environment. *J Assoc Res Otolaryngol* 10:142–149.
5. Strelcyk O, Dau T (2009) Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J Acoust Soc Am* 125:3328–3345.
6. Lőcsei G, et al. (2016) Temporal fine-structure coding and lateralized speech perception in normal-hearing and hearing-impaired listeners. *Trends Hear* 20:2331216516660962.
7. Best V, et al. (2016) On the contribution of target audibility to performance in spatialized speech mixtures. *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing* (Springer, New York), pp 83–91.
8. Kubovy M, Van Valkenburg D (2001) Auditory and visual objects. *Cognition* 80: 97–126.
9. Shinn-Cunningham BG, Best V (2008) Selective attention in normal and impaired hearing. *Trends Amplif* 12:283–299.
10. Shinn-Cunningham BG (2008) Object-based auditory and visual attention. *Trends Cogn Sci* 12:182–186.
11. Bizley JK, Cohen YE (2013) The what, where and how of auditory-object perception. *Nat Rev Neurosci* 14:693–707.
12. Bregman AS (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA).
13. Darwin CJ, Carlyon RP (1995) Auditory grouping. *Hearing*, ed Moore BCJ (Academic, San Diego), pp 387–424.
14. Shinn-Cunningham B (2017) Cortical and sensory causes of individual differences in selective attention ability among listeners with normal hearing thresholds. *J Speech Lang Hear Res* 60:2976–2988.
15. Choi I, Rajaram S, Varghese LA, Shinn-Cunningham BG (2013) Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. *Front Hum Neurosci* 7:115.
16. Woldorff MG, et al. (1993) Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci USA* 90:8722–8726.
17. Hillyard SA, Vogel EK, Luck SJ (1998) Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philos Trans R Soc Lond B Biol Sci* 353:1257–1270.
18. O'Sullivan JA, et al. (2014) Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cereb Cortex* 25:1697–1706.
19. O'Sullivan J, et al. (2017) Neural decoding of attentional selection in multi-speaker environments without access to clean sources. *J Neural Eng* 14:056001.
20. Mirkovic B, Bleichner MG, De Vos M, Debener S (2016) Target speaker detection with concealed EEG around the ear. *Front Neurosci* 10:349.
21. Ding N, Simon JZ (2012) Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J Neurophysiol* 107:78–89.
22. Dai L, Shinn-Cunningham BG (2016) Contributions of sensory coding and attentional control to individual differences in performance in spatial auditory selective attention tasks. *Front Hum Neurosci* 10:530.
23. Hill KT, Miller LM (2010) Auditory attentional control and selection during cocktail party listening. *Cereb Cortex* 20:583–590.
24. Lee AKC, et al. (2013) Auditory selective attention reveals preparatory activity in different cortical regions for selection based on source location and source pitch. *Front Neurosci* 6:190.
25. Choi I, Wang L, Bharadwaj H, Shinn-Cunningham B (2014) Individual differences in attentional modulation of cortical responses correlate with selective attention performance. *Hear Res* 314:10–19.
26. Hopfinger JB, Buonocore MH, Mangun GR (2000) The neural mechanisms of top-down attentional control. *Nat Neurosci* 3:284–291.
27. Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22:751–761.
28. LaBerge D, Auclair L, Sieroff E (2000) Preparatory attention: Experiment and theory. *Conscious Cogn* 9:396–434.
29. Munneke J, Heslenfeld DJ, Usrey WM, Theeuwes J, Mangun GR (2011) Preparatory effects of distractor suppression: Evidence from visual cortex. *PLoS One* 6:e27700.
30. Couperus JW, Mangun GR (2010) Signal enhancement and suppression during visual-spatial selective attention. *Brain Res* 1359:155–177.
31. McEvoy L, Hari R, Imada T, Sams M (1993) Human auditory cortical mechanisms of sound lateralization: II. Interaural time differences at sound onset. *Hear Res* 67: 98–109.
32. Best V, Ozmeral EJ, Kopco N, Shinn-Cunningham BG (2008) Object continuity enhances selective auditory attention. *Proc Natl Acad Sci USA* 105:13174–13178.
33. Ruggles D, Shinn-Cunningham B (2011) Spatial selective auditory attention in the presence of reverberant energy: Individual differences in normal-hearing listeners. *J Assoc Res Otolaryngol* 12:395–405.
34. Cusack R, Deeks J, Aikman G, Carlyon RP (2004) Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *J Exp Psychol Hum Percept Perform* 30:643–656.
35. Hershkowitz RM, Durlach NI (1969) Interaural time and amplitude JND's for a 500 Hz tone. *J Acoust Soc Am* 46:1464–1467.
36. Zwislocki J, Feldman R (1956) Just noticeable differences in dichotic phase. *J Acoust Soc Am* 28:860–864.
37. Smoski WJ, Trahiotis C (1986) Discrimination of interaural temporal disparities by normal-hearing listeners and listeners with high-frequency sensorineural hearing loss. *J Acoust Soc Am* 79:1541–1547.
38. Petersen EB, et al. (2016) Neural tracking of attended versus ignored speech is differentially affected by hearing loss. *J Neurophysiol* 117:18–27.
39. Fiedler L, et al. (2017) Single-channel in-ear-EEG detects the focus of auditory attention to concurrent tone streams and mixed speech. *J Neural Eng* 14:036020.
40. Van Eyndhoven S, Francart T, Bertrand A (2017) EEG-informed attended speaker extraction from recorded speech mixtures with application in neuro-steered hearing prostheses. *IEEE Trans Biomed Eng* 64:1045–1056.
41. Fuglsang SA, Dau T, Hjortkjær J (2017) Noise-robust cortical tracking of attended speech in real-world acoustic scenes. *Neuroimage* 156:435–444.
42. Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
43. Delorme A, Makeig S (2004) EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *J Neurosci Methods* 134:9–21.
44. Baayen R, Davidson D, Bates D (2008) Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang* 59:390–412.