Research Article

Degree of Hearing Loss Affects Bilateral Hearing Aid Benefits in Ecologically Relevant Laboratory Conditions

Todd A. Ricketts,^a Erin M. Picou,^a James Shehorn,^b and Andrew B. Dittberner^c

Purpose: Previous evidence supports benefits of bilateral hearing aids, relative to unilateral hearing aid use, in laboratory environments using audio-only (AO) stimuli and relatively simple tasks. The purpose of this study was to evaluate bilateral hearing aid benefits in ecologically relevant laboratory settings, with and without visual cues. In addition, we evaluated the relationship between bilateral benefit and clinically viable predictive variables.

Method: Participants included 32 adult listeners with hearing loss ranging from mild–moderate to severe–profound. Test conditions varied by hearing aid fitting type (unilateral, bilateral) and modality (AO, audiovisual). We tested participants in complex environments that evaluated the following domains: sentence recognition, word recognition, behavioral listening effort, gross localization, and subjective ratings of spatialization. Signal-to-noise ratio was adjusted to provide similar unilateral speech recognition performance in both modalities and across procedures.

Results: Significant and similar bilateral benefits were measured for both modalities on all tasks except listening effort, where bilateral benefits were not identified in either modality. Predictive variables were related to bilateral benefits in some conditions. With audiovisual stimuli, increasing hearing loss, unaided speech recognition in noise, and unaided subjective spatial ability were significantly correlated with increased benefits for many outcomes. With AO stimuli, these same predictive variables were not significantly correlated with outcomes. No predictive variables were correlated with bilateral benefits for sentence recognition in either modality.

Conclusions: Hearing aid users can expect significant bilateral hearing aid advantages for ecologically relevant, complex laboratory tests. Although future confirmatory work is necessary, these data indicate the presence of vision strengthens the relationship between bilateral benefits and degree of hearing loss.

The decision to fit one or two hearing aids has been a topic of significant clinical and scientific interest for several decades, even for listeners with symmetrical sensorineural hearing loss. Given the additional cost, inconvenience, and potential stigma of fitting hearing aids bilaterally, it is important for practitioners to be able to make evidence-based recommendations for patients regarding expected benefits. Many laboratory studies have demonstrated significantly better performance when individuals use two hearing aids compared to one, referred to herein as bilateral benefit (Freyaldenhoven, Plyler, Thelin,

& Burchfield, 2006; Hawkins & Yacullo, 1984; Köbler & Rosenhall, 2002; Markle & Aber, 1958).

Despite evidence of significant bilateral benefits, the number of patients who ultimately choose a bilateral fitting is not consistently high. Researchers who have examined whether listeners prefer one versus two hearing aids have reported that preference for two hearing aids ranges approximately from 30% to 55% in field studies (Cox, Schwartz, Noe, & Alexander, 2011; Day, Browning, & Gatehouse, 1988; Erdman & Sedge, 1981; Schreurs & Olsen, 1985; Stephens et al., 1991; Vaughan-Jones, Padgham, Christmas, Irwin, & Doig, 1993) and approximately from 70% to 95% in retrospective studies (Bertoli, Bodmer, & Probst, 2010; Boymans, Goverts, Kramer, Festen, & Dreschler, 2009; Chung & Stephens, 1986; Dillon, Birtles, & Lovegrove, 1999; Köbler, Rosenhall, & Hansson, 2001). When listeners are fitted with their preferred fitting type (unilateral vs. bilateral), hearing aid outcomes have generally been shown to be similar on indices of use, satisfaction, benefit, and

^aDepartment of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN

^bHeuser Hearing Institute, Louisville, KY

^cGN Hearing, Chicago, IL

Correspondence to Todd A. Ricketts: todd.a.ricketts@vumc.org

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Christian Stilp

Received January 11, 2019

Revision received May 13, 2019

Accepted July 2, 2019

https://doi.org/10.1044/2019_JSLHR-H-19-0013

Disclosure: The first author received an investigator-initiated grant from GN ReSound to support this work. The fourth author works for the company that was the primary funding source, but no products are promoted in this research. There are no other conflicts of interest.

residual handicap (Boymans et al., 2009; Stephens et al., 1991; T. C. Walden & Walden, 2004), although some studies have shown improved outcomes on these dimensions with bilateral fittings (Bertoli et al., 2010; Cox et al., 2011).

One potential explanation for the apparent discrepancy between demonstrated laboratory benefits and patient preferences is the limited ecological relevance of previous laboratory studies. Several investigators have suggested that accurate descriptions of hearing aid benefits should include methods with a focus on ecologic validity (Brody, Wu, & Stangl, 2018; Day et al., 1988; Gatehouse, Elberling, & Naylor, 1999; Miller et al., 2017), that is, evaluations that more directly emulate complex, real-world listening environments. Such evaluations of hearing aid benefits in prototypical listening situations that differ in the demands placed on the listener have been advocated for previously, because they are expected to better reflect benefits realized by patients in real-world settings (Cox, Alexander, & Gilmore, 1987b; B. E. Walden, 1997; Wu et al., 2018). Importantly, assessment in real-world settings presents its own set of challenges, including less control, poor test–retest reliability, and subjective reporting biases, all of which can reduce sensitivity to effects (Guyatt, Oxman, Kunz, Brozek, et al., 2011; Guyatt, Oxman, Kunz, Woodcock, et al., 2011; Guyatt, Oxman, Vist, et al., 2011).

There is the possibility that testing under conditions of increased ecological relevance may decrease or eliminate bilateral benefits that were previously identified in simple laboratory environments (Silberer, Bentler, & Wu, 2015; Wu & Bentler, 2010). A lack of bilateral benefits under ecologically valid laboratory conditions could help explain the apparent discrepancy between laboratory and field studies. One way to increase the ecological relevance of laboratory testing is to examine bilateral benefits across multiple domains simultaneously, as typical listening situations are often complex, requiring speech recognition, cognition, and spatialization.

Bilateral Benefits in Multiple Listening Domains

Ecologically relevant domains with the potential to be sensitive to bilateral fittings include speech recognition, listening effort, spatialization performance, and ratings of sound quality. Each will be discussed in turn. For all domains, it is important to distinguish the term *bilateral benefit* from the term *binaural benefit*, which is used to describe the benefit of having two ears compared to one. While wearing hearing aids, it is common for some binaural information to be audible, even when fitted with one instrument. This is most likely for individuals with sloping hearing loss and significant venting. Consequently, bilateral benefits are not always consistent with, or of the same magnitude as, binaural benefits. In addition, the relationship between binaural and bilateral benefits is affected by degree of hearing loss, as we describe below.

First, bilateral speech recognition benefits are most commonly reported in studies that employed spatially separated speech and noise sources (Freyaldenhoven et al., 2006;

Hawkins & Yacullo, 1984; Köbler & Rosenhall, 2002; Markle & Aber, 1958). When speech and noise are collocated, a smaller percentage of listeners exhibit benefits (B. E. Walden & Walden, 2005). These findings are consistent with the work demonstrating the benefits of binaural listening for speech understanding with spatially separated noise (Best, Mason, & Kidd, 2011). In addition, the majority of these studies used a single target source location presented in an environment with little reverberation. The potential for bilateral speech recognition benefits in more complex listening environments is not yet well understood.

Second, several researchers have reported subjective benefits of listening effort with bilateral fittings (Most, Adi-Bensaid, Shpak, Sharkiya, & Luntz, 2012; Noble, 2006; Noble & Gatehouse, 2006; van Schoonhoven et al., 2016). Given the apparent dissociation between objective and subjective measures of listening effort (Feuerstein, 1992; Hicks & Tharpe, 2002; Moore & Picou, 2018), it is important to also evaluate potential bilateral benefits for listening effort objectively. On the basis of the findings that hearing aids reduce listening effort (Downs, 1982; Picou, Ricketts, & Hornsby, 2013), we also expect bilateral hearing aid benefits for listening effort, although they have not been demonstrated previously in the literature.

Third, bilateral benefits for localization have been reported for some conditions. Specifically, bilateral fittings generally allow for better localization than unilateral fittings, as indicated by subjective reports of improved localization (Boymans et al., 2009; Boymans, Goverts, Kramer, Festen, & Dreschler, 2008; Köbler et al., 2001; Stephens et al., 1991; Vaughan-Jones et al., 1993) and laboratory tests of auditory localization in the horizontal plane (Boymans et al., 2008; Byrne, Noble, & LePage, 1992; Köbler & Rosenhall, 2002). As with speech recognition, these previous studies of localization were conducted in controlled laboratory settings with relatively simple tasks.

However, real-world localization tasks are often not simple. For example, consider a hearing aid wearer entering a crowded, noisy gathering in a moderately reverberant room and someone begins speaking to them. Optimal communication may require that the listener quickly and accurately locate the talker, obtain visual cues, recognize the speech, and engage cognitively with the spoken message. This task, although requiring basic localization skills, is more complex. On the basis of the localization task only, we would expect bilateral benefits in the complex scenario with overlapping tasks, but benefits have not been reported.

Finally, several studies have also reported better subjective ratings of sound quality with bilateral fittings on dimensions such as clarity, loudness, and balance (Balfour & Hawkins, 1992; McKenzie & Rice, 1990; Naidoo & Hawkins, 1997). Similarly, in real-world trials, subjective preferences also generally tend to favor bilateral fittings (Boymans et al., 2008; Cox et al., 2011; Erdman & Sedge, 1981; Köbler et al., 2001; Punch, Jenison, Allan, & Durrant, 1991). For example, Köbler et al. (2001) surveyed experienced hearing aid users. The majority of respondents reported better speech recognition and better overall sound

quality with bilateral hearing aids. In addition, most respondents reported that bilateral hearing aid use was beneficial when attending a lecture, in group conversations, while listening to music, and while watching television. Exceptions to positive sound quality ratings with two hearing aids are on the dimension of loudness, with some patients reporting that bilateral fittings are less comfortable or are too loud (Boymans et al., 2008; Cox et al., 2011).

Audiovisual Presentations

Another method of increasing ecological relevance is to evaluate bilateral benefits with audiovisual (AV) stimuli. Although audio-only (AO) situations lend themselves well to laboratory study, they are not representative of many real-world situations during which the talker's face is often visible. Each of the aforementioned domains can also be affected by the presence of visual cues. For speech recognition, visual cues have been shown to improve performance over a wide range of signal-to-noise ratios (SNRs; Erber, 1975; Grant & Seitz, 1998; O'Neill, 1954; Sumby & Pollack, 1954), particularly in difficult listening situations, such as when usable audibility is limited (Bernstein & Grant, 2009; Helfer & Freyman, 2005) or when context is limited (Grant, Walden, & Seitz, 1998). Visual cues have also been shown to interact with hearing aid benefit. For example, the presence of visual cues has been shown to decrease directional benefit in some listening situations (Wu & Bentler, 2010), as well as decrease the audible bandwidth necessary for accurate speech recognition (Silberer et al., 2015). The interaction between bilateral benefits and visual cues is less clear. While at least one study has demonstrated significant bilateral speech recognition benefits using an AV task (Day et al., 1988), this study was limited to individuals with severe hearing loss.

The addition of visual cues is also expected to improve auditory localization, since once the target is in the field of vision, its location can often easily be identified. Whereas most existing literature reports on the degree to which visual cues bias auditory cues when the two modalities are in conflict (e.g., Jackson, 1953; Welch & Warren, 1980), few studies report on the effect of adding visual cues to localization tasks. Intuitively, the addition of congruent visual cues could substantially facilitate auditory localization, particularly for signals of longer duration. Indeed, at least one previous investigation, which explored word recognition and gross localization of a speech signal, revealed better localization performance when congruent visual cues were available than when stimuli were auditory only (Picou, Aspell, & Ricketts, 2014). Importantly, the task used in this study was not a simple localization task. Instead, it was designed as an ecologically relevant measure that required listeners to quickly locate a speech source from four possible locations and repeat a string of words presented from that source.

The expected effects of visual cues on listening effort are more complicated. When the addition of visual cues leads to improved speech recognition, concurrent listening effort also improves (Fraser, Gagné, Alepins, & Dubois, 2010). Conversely, there may be cognitive cost for some

listeners to integrate auditory and visual cues if speech recognition performance is matched in AO and AV conditions (Fraser et al., 2010), notably for listeners who are not skilled lip readers (Picou, Ricketts, & Hornsby, 2011).

Individual Variability in Bilateral Benefits

In addition to the ecological relevance of test conditions, lower-than-expected patient preference for bilateral hearing aids might be attributed to individual patient factors. The fact that many patients prefer unilateral fittings is also commonly attributed to factors such as price, cosmetics, and inconvenience (Boymans et al., 2009; Cox et al., 2011). Several studies have attempted to explain the variability in bilateral benefits using predictive variables. In general, audiometric configuration, degree of asymmetry, and demographic data (e.g., age, gender) have not been shown to be significantly related to preference for bilateral hearing aids (Boymans et al., 2009; Day et al., 1988; Köbler et al., 2001; Stephens et al., 1991; Vaughan-Jones et al., 1993).

Patient expectations regarding limited bilateral benefits is also a significant factor for bilateral preference (Boymans et al., 2009; Cox et al., 2011). Indeed, a hallmark of published bilateral benefit data is high individual variability. That is, whereas some listeners demonstrate significant bilateral benefits across multiple domains, other participants demonstrate limited or no significant benefits. Some listener traits have been identified as noncontributing factors. For example, results from previous investigations suggest that bilateral speech recognition benefit could not be predicted based on tests of binaural processing using headphones (Boymans et al., 2008) or based on degree of loudness summation, hearing handicap, or personality (Cox et al., 2011). Conversely, two specific listener traits have often been identified as successful predictors of bilateral benefit: binaural interference for speech recognition and degree of hearing loss.

Binaural interference for speech recognition refers to a dichotic deficit wherein binaural performance with two ears is measurably worse than monaural performance during a situation where the reverse would be expected. That is, the additional auditory information from the second ear interferes with performance. Although binaural interference is relatively rare (prevalence estimate of between 5% and 18% of the general population; Allen, Schwab, Cranford, & Carpenter, 2000; Mussoi & Bentler, 2017), several authors have reported that binaural interference can be a predictor for unsuccessful bilateral hearing aid use (Carter, Noe, & Wilson, 2001; Chmiel, Jerger, Murphy, Pirozzolo, & Tooley-Young, 1997; Jerger, Silman, Lew, & Chmiel, 1993; Köbler, Lindblad, Olofsson, & Hagerman, 2010). The degree of hearing loss can have differential effects on binaural versus bilateral benefits. Specifically, binaural benefits often decrease with increasing degree of hearing loss (e.g., Durlach, Thompson, & Colburn, 1981). However, bilateral benefits can increase with degree of hearing loss because of changes in audibility. Specifically, decreasing hearing loss increases audibility for sounds in the unaided ear. Therefore, listeners

fitted with only one hearing aid and usable unaided hearing in their better ear may still be able to make use of the unaided auditory information. In these cases, the addition of the second hearing aid does not change the listening situation from monaural to binaural; rather, the second hearing aid simply improves the sensation level and, in some cases, the audible bandwidth, in the second ear. Therefore, it is expected that the second hearing aid would be more beneficial for listeners with more hearing loss. Indeed, several investigators have reported greater and more consistent bilateral benefits for listeners with greater degrees of hearing loss on measures of speech recognition (Festen & Plomp, 1986; Hedgecock & Sheets, 1958; Jerger, Carhart, & Dirks, 1961; McArdle, Killion, Mennite, & Chisolm, 2012), localization (Byrne et al., 1992), preference (Chung & Stephens, 1986), and subjective ratings of benefit (Boymans et al., 2008; Noble, 2006; van Schoonhoven et al., 2016). More recently, van Schoonhoven et al. (2016) demonstrated that the magnitude of bilateral benefits generally increases with increasing hearing loss in multiple domains. Specifically, although listeners with moderate-to-severe hearing impairment demonstrated bilateral benefits for speech recognition, subjective listening effort, and localization, significant bilateral benefits were only noted in the domain of subjective listening effort in listeners with mild hearing impairment (pure-tone averages (PTAs) < 40 dB HL).

Purpose

The purpose of this study was to evaluate bilateral hearing aid benefit in ecologically relevant conditions, specifically across multiple auditory domains and with AO or AV stimuli. The prototypical listening situations of interest included face-to-face communication, in addition to talker identification and speech recognition when multiple sources are present. The specific domains of interest were speech recognition, gross localization, behavioral listening effort, and subjective spatialization. Because the addition of visual cues reduces a listener's reliance on auditory cues, we expect that bilateral benefits will be reduced in AV situations, particularly for speech recognition and localization. In addition to adding visual cues, tasks were completed in moderate reverberation, with noise surrounding the listener, and in some cases, several domains were evaluated simultaneously. If bilateral benefits are reduced under the ecologically relevant conditions, it might help explain the apparent discrepancy between bilateral benefits in the laboratory and patient preferences. Since the magnitude of bilateral benefit in the presence of visual cues had not previously been systematically explored, possible predictive relationships based on clinically viable measures, including age, unaided speech recognition in noise, unaided spatial release from masking (SRM), and unaided subjective spatial ability, were also considered. Given the relative rarity of binaural interference for speech recognition, it was not assessed as a predictive factor in the current experiment as a very large sample population would be required.

Method

Participants and Instruments

Thirty-two native English-speaking adults aged 40–85 years ($M = 67.9$ years) participated in this study. Recruitment targeted participants with hearing loss ranging from mild sloping to moderate in the high frequencies (1000–4000 Hz) to listeners with flat severe configurations. Figure 1 displays the mean and individual better ear audiometric data for all participants. Hearing losses were symmetrical for all participants, as evidenced by interaural differences equal to or less than 15 dB at any one audiometric frequency and equal to or less than 10 dB at three consecutive frequencies. Participants had sensorineural hearing loss as evidenced by normal acoustic immittance findings and air–bone gaps < 15 dB (500–4000 Hz). Participants denied history of neurogenic, cognitive, or otologic disorders. All testing was conducted with approval from Vanderbilt University's Institutional Review Board. Participants were compensated for their time.

All participants were fitted with a pair of hearing aids for the purposes of laboratory testing. Listeners with less hearing loss were fitted with commercially available, multichannel compression, behind-the-ear hearing aids with Comply ear tips using clinically appropriate venting. While these foam ear tips are noncustom, venting is applied via various sizes of “trough venting” channels manufactured into the side of the ear tips. This system limits the maximum equivalent vent size to approximately 2 mm. Listeners with greater degrees of hearing loss were fitted with a commercially available, power behind-the-ear hearing aid using the same multichannel compression processing coupled to an unvented Comply ear tip.

For all listeners, the hearing aids were programmed with a single program that activated an omnidirectional microphone setting and digital feedback reduction. All other advanced features were disabled (i.e., directional microphones, digital noise reduction, speech enhancement, and

Figure 1. Overall mean and individual better ear audiometric data for study participants.

impulse noise reduction). For each participant, the hearing aid gain was adjusted to match National Acoustic Laboratories' Nonlinear 1 prescriptive targets (Byrne, Dillon, Ching, Katsch, & Keidser, 2001) and verified with probe microphone measures. Verification was accomplished with an AudioScan Verifit, which uses a recorded speech signal (i.e., carrot passage) presented from a built-in loudspeaker at 0°. Hearing aid output in both ears was verified to be within ± 4 dB of prescriptive target for input levels of 55, 65, and 75 dB SPL from 250 to 4000 Hz.

Measures of Predictive Variables

All participants completed two sets of tests: potential predictive measures (unaided) and main outcome measures (aided testing with one or two hearing aids). All predictive measures were completed in a custom sound booth (4 m $\times$ 4.3 m $\times$ 2.7 m). During an initial laboratory visit, following informed consent procedures, participants underwent pure-tone audiometric testing and acoustic immittance testing in accordance with standard clinical procedures. Also, each participant's subjective spatial abilities, unaided speech recognition in noise, and unaided SRM abilities were evaluated.

Subjective spatial abilities were evaluated via the 12-item version of the Speech, Spatial and Qualities of Hearing Scale (SSQ-12). This scale of subjective spatial abilities has been shown to provide similar outcomes to a 49-item version in a large clinical research sample (Noble, Jensen, Naylor, Bhullar, & Akeroyd, 2013). Participants rate each item on an 11-point Likert scale (0–10 rating points), with a lower rating indicating more difficulty. A composite score was created by averaging all 12 items; the mean SSQ-12 score was 5.2 (range: 1.5–8.7). This mean value is quite similar to that reported previously for elderly listeners with moderate hearing loss (5.5) on the spatial subscale of the SSQ (Gatehouse & Noble, 2004).

Unaided speech recognition testing was completed via the Bamford-Kowal-Bench Speech-in-Noise Test (Killion, Niquette, Revit, & Skinner, 2001) presented diotically to both ears through headphones (Etymotic ER-3A). The BKB-SIN test yields an estimated SNR for 50%-correct recognition of key words in sentences (SNR-50) when the speech is presented at a high presentation level (83 dB SPL). The SNR of the Bamford-Kowal-Bench Speech-in-Noise Test decreases in 3-dB steps from +21 to –6 dB. Each list pair consists of two lists, each with 10 sentences. Performance was evaluated following test instructions using two list pairs. The mean unaided SNR-50 score was 3.6 dB (range: –3 to 19 dB).

SRM was measured using the Hearing in Noise Test (HINT; S. Soli & Nilsson, 1994) presented via loudspeakers (Bowers & Wilkins 685 S2). The HINT is an adaptive SNR test where the SNR required to achieve 50% speech understanding is estimated by varying the level of speech in spectrally matched, steady-state background noise with a fixed level (65 dB SPL). The presentation level of the sentences is varied in 4-dB step sizes for four reversals and then in 2-dB

step sizes for the remainder of the list. The noise level is subtracted from the mean presentation level during the 2-dB step-size reversals to estimate the 50% performance SNR. Two lists of 10 HINT sentences were presented consecutively in each of two speech and noise configurations, collocated and separated. In the collocated condition, speech and noise originated from a loudspeaker located 1 m directly in front of a participant. In the separated condition, the speech signal originated 1 m directly in front of a participant, and the noise originated from a loudspeaker positioned at 90° (on the participant's left side). The difference between the SNR in the collocated and separated conditions was used as an indication of a participant's SRM. The mean SRM score was 4.1 dB (range: –3.1 to 11.5 dB). Only two participants demonstrated negative SRM scores, indicating similar or better performance in the collocated condition when compared to the spatially separated condition.

Main Outcomes: Test Environment and Apparatus

All main outcome measures were completed in a moderately reverberant environment ($T_{30} = 475$ ms; 5.5 m $\times$ 6.5 m $\times$ 2.25 m). Floor carpet, ceiling acoustic blankets (Sound Spotter 124 4 $\times$ 4), and wall acoustic blankets (Sound Spotter 124 4 $\times$ 8) were hung to control reverberation. All target stimuli were presented from one loudspeaker (Tannoy System 600A) placed 1.25 m directly in front of the participant or from one of four loudspeakers (Tannoy System 600A) placed at $\pm 45^\circ$ and $\pm 60^\circ$, 1.25 m from the participant. Each target loudspeaker had a Dell 24-in. widescreen computer monitor placed directly on top for delivery of visual stimuli during AV conditions. During AO conditions, the monitor(s) was powered off.

Competing noise stimuli were played from a computer via sound editing software (Adobe Audition CC 5.5) to a multichannel soundcard (Echo Layla 3G), then to a multichannel amplifier (Crown CTS 8200), and to four noise loudspeakers (Tannoy System 600). The noise loudspeakers were 3.5 m from the participant and located at 45°, 135°, 225°, and 315°. The same four-talker babble was used for all testing. The four talkers were recordings of female talkers reading sentences from the Connected Speech Test (CST; Cox, Alexander, & Gilmore, 1987a; Cox, Alexander, Gilmore, & Pusakulich, 1989). All sentences were presented at the same root-mean-square level. A single unique talker was presented from each of the four loudspeakers. The specific location of a given talker moved such that the talker's voice originated from all four loudspeakers over time. The System 600 and 600A test loudspeakers have a $Q = 6$. Consequently, the calculated critical distance was approximately 1.4 m (Peutz, 1971). Thus, the listener was always seated within the critical distance of the speech stimuli loudspeaker. The constant distance of 1.25 m was selected to be representative of a distance for face-to-face communication near the boundary between common distances found for conversation between friends and those greater distances common in social and business situations

(Hall, 1966). The same background noise routing was used for all tasks.

Main Outcomes: Stimuli and Procedure

Following unaided testing, each participant was fitted bilaterally with hearing aids, and the remaining tests were performed, which included four experimental tasks exploring the domains of sentence recognition, word recognition, gross localization, subjective ratings of spatialization, and behavioral listening effort. These four tasks are described in detail below and summarized in Table 1. For each aided domain, testing was completed in two modalities (AO, AV) and two hearing aid fitting types (unilateral, bilateral). In addition to AV evaluation, some of the experimental tasks were designed to evaluate performance in more than one domain simultaneously in an attempt to further improve ecological relevance. In the AV conditions, a video recording of the talker's face was visible on a screen. In the unilateral conditions, the hearing aid from the non-test ear was removed, but the ear was not plugged. The unaided ear for the unilateral test conditions was counterbalanced across all participants. In addition, the order of conditions within each domain was counterbalanced to avoid learning and fatigue effects. For each participant, testing occurred over the course of 2–4 testing days.

It has previously been demonstrated that real-world SNRs are commonly quite positive (from +2 to 14 dB), with negative SNRs rarely occurring (Pearsons, Bennett, & Fidell, 1977; Smeds, Wolters, & Rung, 2015; Wu et al., 2018). However, while ecological relevance was a test goal, it was also of interest to choose conditions that resulted in similar unilateral performance across test conditions and individuals given the known influence of the psychometric function on the magnitude of visual benefit and to avoid ceiling and floor effects. Given that previous experience with the test materials has shown that those with hearing loss in the severe range were expected to have ceiling performance

well below 100%, a relatively low performance level was targeted (approximately 50%). Pilot testing was first completed to select SNRs and speech presentation levels that led to approximately 50% speech recognition performance for each unilateral test condition as a function of degree of hearing loss. Then, individual participants were tested at one of three sets of SNRs for each test, depending on their degree of hearing loss and performance on practice lists, as specified below. While allowing for some natural variability in performance, this method helped ensure that those with greater and lesser degrees of hearing loss were tested over similar ranges of the psychometric function for each test procedure. Consequently, listeners with greater hearing loss were evaluated in SNRs that are more common in the real world (e.g., from +7 dB to Quiet), whereas listeners with less hearing loss were evaluated in more difficult SNRs (e.g., from –4 to +7 dB), some of which are less common in the real world.

Sentence Recognition in Noise

It is common to assess speech recognition with limited context, since cognitive processing abilities can influence outcomes. However, conversational partners often know the topic being discussed in one-on-one conversations. To capture performance in this listening situation, the CST (Cox et al., 1987a, 1989) was used to assess sentence recognition in noise. The CST also represents one of the few speech recognition tests with published normative data for AO and AV presentations. The CST contains 24 passage pairs; each passage contains 10 sentences about a common topic (e.g., giraffes). Consistent with test instructions, researchers informed the participants of the target topic before the presentation of each passage. Scoring was based on 50 key words presented in each passage pair. Two passage pairs were used for each condition, for a total of 100 scoring key words. All conditions were tested once before the conditions were repeated. Speech stimuli were played from a commercially available DVD player

Table 1. The main experimental outcomes within the four experimental tasks.

Task name	Skill assessed	Outcome measure(s)	AO SNRs (dB)	AV SNRs (dB)	Speech level (dBA)
Connected Speech Test	Single-talker sentence recognition with known topic	Sentence recognition performance (rau)	+2	–4	58
			+11	+7	65
			Quiet	+11	65
Spatial Test Requiring Effortful Speech Recognition (STRESR)	Localization speed, localization accuracy, word recognition with variable location	1. Localization response time (ms)	+7	+3	65
		2. Localization accuracy (%)	+11	+7	65
		3. Word recognition with variable talker location (rau)	Quiet	+11	65
Dual-task paradigm	Listening effort, word recognition with fixed location	1. Word recognition performance with fixed talker location (rau)	+7	+3	65
			+11	+7	65
		2. Response times (ms)	Quiet	+11	65
Subjective spatial ratings	Subjective spatial impression following the STRESR	Mean score on subjective rating scale (seven questions)	+7	+3	65
			+11	+7	65
			Quiet	+11	65

Note. The presentation levels and signal-to-noise ratios (SNRs) tested within groups of listeners within each outcome are also listed. AO = audio-only; AV = audiovisual.

(Pioneer DV-563A) routed to an attenuator for level control (TDT System 2 PA5) and then to a single loudspeaker directly in front of a participant.

Prior to data collection, all participants practiced using passage pairs that were not used for testing. For participants with PTA thresholds less than 55 dB HL ($N = 16$), the overall level of speech was 58 dBA, whereas the levels of background noise were 56 dBA (AO, +2 dB SNR) and 62 dBA (AV, -4 dB SNR). For participants with PTA thresholds equal to or greater than 55 dB HL ($N = 16$), there was significant variability in participants' abilities to understand speech; hence, the level of background noise was varied to avoid floor effects. The level of speech was always 65 dBA, and the overall level of background noise was chosen based on the participant's performance during practice. Six participants were tested with background noise levels of 55 dBA (AO, +11 SNR) and 58 dBA (AV, +7 SNR); 10 participants were tested in quiet (AO) and with a background noise level of 54 dBA (AV, +11 SNR).

Spatial Test Requiring Effortful Speech Recognition (Gross Localization and Word Recognition With Variable Location)

The Spatial Test Requiring Effortful Speech Recognition (STRESR) was used to evaluate gross localization and word recognition simultaneously to assess some of the skills needed when listening in an environment with multiple potential talkers with different locations. Details of STRESR development are reported in Picou et al. (2014). During testing, five words were presented from one of four loudspeakers located at $\pm 45^\circ$ and $\pm 60^\circ$. The participant's task was to indicate as quickly as possible, using a USB keypad, which loudspeaker was presenting the stimuli. After all five words were presented, the participant's task was to repeat as many of the five words as possible. Since this task requires listeners to recognize and repeat a string of five unrelated words, word recognition performance has the potential to be affected by memory. This task provides three outcomes: (a) gross localization accuracy as measured by percent-correct speaker location identification, (b) gross localization speed as measured by reaction time, and (c) word recognition. Four different word lists (120 words in each) were used, two for AO conditions and two for AV conditions. The order of lists and conditions was counterbalanced across participants. Prior to data collection, participants practiced the task at least twice.

The speech stimuli used in the gross localization task were monosyllabic words recorded by a female talker. In each list of 120 words, six sets of five words were presented from each loudspeaker. The speech stimuli were presented using custom programming via Presentation software (Neurobehavioral Systems) and routed to an electronic switch (Extron AV). Then, the speech stimuli were routed to one of four loudspeaker-and-monitor combinations. The other three loudspeakers were silent, and the other three monitors displayed the same talker's face saying nontarget words without the auditory signals. During AV conditions, no faces were visible. For participants

with PTA thresholds less than 55 dB HL, the overall level of background noise was 58 dBA (AV conditions; +7 SNR) or 62 dBA (AV conditions; +3 SNR). For participants with PTA thresholds equal to or greater than 55 dB HL, the background noise was of the same level as was used for speech recognition testing (+11/+7 or Quiet/+11 for AV/AV conditions).

Subjective Spatial Ratings

Immediately following each STRESR condition, participants were asked to rate their subjective spatial impression related to listening on a 10-point Likert scale. The specific questions and anchors used for the end points included the following: (a) How easy was it for you to monitor what was going on around you? (*not at all easy* to *very easy*); (b) How easy was it for you to tell where sounds were coming from? (*not at all easy* to *very easy*); (c) How easy was it to find the talker? (*not at all easy* to *very easy*); (d) Did you have the impression of the sounds being exactly where you expected? (*not at all* to *perfectly*); (e) Could you easily ignore other sounds when trying to listen to the main sounds? (*could not easily ignore* to *could easily ignore*); (f) Could you tell where the talker was as soon as they started speaking? (*not at all* to *perfectly*); and (g) Could you tell right away whether the talker was on your left or your right, without having to look? (*not at all* to *perfectly*). The average responses from the seven questions were calculated to arrive at a single subjective spatial inventory score for each of the four test conditions.

Dual-Task Paradigm (Listening Effort and Word Recognition With Fixed Location)

A dual-task paradigm was used to objectively evaluate listening effort and word recognition. The primary task was monosyllable word recognition, and the secondary task was to press a button on a USB keypad in response to a visual probe. The stimuli and procedures used during listening effort testing have been reported elsewhere (Picou et al., 2014). During testing, a rectangle was presented 50 ms immediately before word presentation. The rectangle, visible for 125 ms, was either red (probe condition) or white (nonprobe condition). Participants were instructed to press a red button on a USB keypad as quickly as possible whenever a red rectangle appeared, but not to press any button if the rectangle was white. Median reaction time to the probe stimuli was taken as an objective measure of listening effort. Probe trial presentations were randomized within constraints, such that a probe occurred once during every block of four trials (i.e., words), but randomly within each block.

The same monosyllabic words and background noise that were used for gross localization were also used to evaluate listening effort. Instead of being arranged into groups of five words to be presented from one of four loudspeakers, the words were presented in isolation from a loudspeaker directly in front of a listener. Speech stimuli were presented using custom programming via Presentation software (Neurobehavioral Systems Version 12.0). The same speech

and noise levels as during the gross localization testing were used for listening effort testing (+7/+3 SNRs for the moderate group; +11/+7 or Quiet/+11 for the severe group for the AO/AV conditions).

The participant's task was to repeat the word, regardless of whether a probe or a nonprobe trial was presented. After the participant repeated the word, the investigator pushed a button on another keyboard outside of the test booth to begin the next trial. The investigator also scored the participant's verbal response. For AV conditions, a video of the talker's face was visible on a computer monitor. For AO conditions, only a still picture of the talker's face was visible on the monitor. Prior to data collection, participants practiced the listening effort task first by performing only the secondary task, then by practicing both primary and secondary tasks simultaneously in quiet, and, finally, by practicing both primary and secondary tasks simultaneously in noise. Reaction times from participants were accepted as correct only if they occurred during a probe trial.

Data Analysis

Main outcome measures included seven scores for each participant, obtained from four tasks: sentence recognition task (sentence recognition performance score), STRESR task (word recognition with variable location score, localization speed score, and localization accuracy score), listening effort task (word recognition with fixed location score and secondary task response times score), and subjective rating of spatialization (subjective rating score). Before analysis, all sentence recognition and word recognition scores were converted into rau to normalize the variance near the extremes (Studebaker, 1985). Data in this study were normally distributed and met all assumptions necessary for a multivariate, within-subject analysis of variance.

Each of the seven scores was analyzed separately using a repeated-measures analysis of variance (RM-ANOVA) with two within-subject factors, modality (AO, AV) and hearing aid fitting type (unilateral, bilateral). It should be noted that any significant main effects of modality resulted from experimental design since different SNRs were chosen for the AO and AV presentations. Therefore, the only statistical outcomes of interest for this analysis were significant interactions between fitting type and modality and significant main effects of fitting type. Significant interactions were explored using multiple RM-ANOVAs with a single factor (fitting) for each modality (AO or AV). Significant main effects were explored using pairwise comparisons and controlling for familywise error rate with Bonferroni corrections (Dunn, 1961).

To analyze the potential relationship between bilateral benefit and predictive variables, bilateral benefit scores were calculated for each of the seven scores in both modalities. The potentially predictive variables (age, PTA, SRM, SRT-50, SSQ-12) were entered into a partial correlation analysis with the calculated benefit scores. Eight correlation analyses were conducted, one for each of the four tasks. Each analysis included the outcomes from the task (e.g., word

recognition, response times, localization accuracy, and subjective ratings in the case of the STRESR) and four predictive variables (PTA, SNR-50, SSQ, SRM). Age and the corresponding unilateral word or sentence recognition performance score were entered as control variables.

Results

Main Outcome Measures

Sentence Recognition

Figure 2 displays sentence recognition performance for unilateral and bilateral conditions with AO (left panel) and AV (right panel). Analysis of performance scores revealed a significant main effect of hearing aid fitting, $F(1, 31) = 19.57$, $p < .001$, $\eta_p^2 = .39$. The main effect of modality and the Modality $\times$ Hearing Aid interaction were not significant ($p > .20$). Sentence recognition performance was significantly better with two hearing aids compared to one hearing aid, $M = 9.13$ rau, 95% CI [4.92, 13.34].

STRESR

Figure 3 displays the word recognition (top panels), localization speed (middle panels), and localization accuracy (bottom panels) obtained during the STRESR task. Analysis of word recognition scores revealed a significant main effect of hearing aid fitting, $F(1, 31) = 6.57$, $p < .05$, $\eta_p^2 = .18$. The main effect of modality and the Modality $\times$ Hearing Aid interaction were not significant ($p > .10$). These data reveal word recognition performance with a variable talker location was better with bilateral than with unilateral hearing aid fitting, $M = 4.75$ rau, 95% CI [0.97, 8.53]. Analysis of localization response times revealed a significant effect of hearing aid fitting, $F(1, 31) = 31.80$, $p < .001$, $\eta_p^2 = .51$. The main effect of modality and the Modality $\times$ Hearing Aid interaction were not significant ($p > .10$). Localization was faster with two hearing aids compared to one, $M = 1026.30$ ms, 95% CI [655.13, 1297.47].

Analysis of localization accuracy revealed significant effects of hearing aid fitting, $F(1, 31) = 26.32$, $p < .0001$, $\eta_p^2 = .46$, and modality, $F(1, 31) = 64.28$, $p < .001$, $\eta_p^2 = .66$, and a Modality $\times$ Hearing Aid interaction, $F(1, 31) = 15.07$, $p < .01$, $\eta_p^2 = .33$. To explore the significant interaction, follow-up RM-ANOVAs were completed with the AO and AV stimuli separately with a single factor (hearing aid fitting). With AO stimuli, there was a significant effect of hearing aid fitting, $F(1, 31) = 9.63$, $p < .001$, $\eta_p^2 = .49$; localization accuracy was better with two hearing aids, $M = 22.66$ percentage points, 95% CI [14.17, 31.15]. With AV stimuli, there was also a significant effect of hearing aid fitting, $F(1, 31) = 6.81$, $p < .05$, $\eta_p^2 = .18$; localization accuracy was better with two hearing aids, $M = 7.16$ percentage points, 95% CI [1.56, 12.76].

Listening Effort

Figure 4 displays the word recognition (top panels) and response times (bottom panels) measured during the

Figure 2. Scores for the sentence recognition task with audio-only (left panel) and audiovisual (right panel) modalities. Open symbols and dashed lines indicate performance with unilateral hearing aid; filled symbols and solid lines indicate performance with bilateral hearing aids. Data are presented as a function of better ear, pure-tone average (dB HL), calculated based on thresholds at 500, 1000, and 2000 Hz.

listening effort task. Analysis of word recognition scores (with a fixed talker location) revealed significant effects of modality, $F(1, 31) = 173.65$, $p < .0001$, $\eta_p^2 = .85$, and hearing aid fitting, $F(1, 31) = 17.51$, $p < .0001$, $\eta_p^2 = .36$. The interaction was not significant. Word recognition was better with AV stimuli, $M = 16.76$ rau, 95% CI [14.17, 19.36], and with two hearing aids, $M = 4.67$ rau, 95% CI [1.56, 12.76]. Analysis of response times during the secondary task revealed nonsignificant effects of modality, hearing aid fitting, and the Modality $\times$ Hearing Aid Fitting interaction. The mean bilateral hearing aid benefit was small and not significant at 4.50 ms, 95% CI [3.6, 12.6].

Subjective Ratings

Figure 5 displays the subjective ratings. Analysis of subjective ratings revealed significant effects of modality, $F(1, 31) = 8.54$, $p < .01$, $\eta_p^2 = .22$, and hearing aid fitting, $F(1, 31) = 32.36$, $p < .0001$, $\eta_p^2 = .51$. Perceived spatialization was better with AV stimuli, $M = 0.68$ points, 95% CI [0.21, 1.15], and with two hearing aids, $M = 1.71$, 95% CI [1.1, 2.33]. The interaction was not significant.

Predictive Variables

The relationships between PTA and the other predictive measures are displayed in Figure 6, including the SSQ-12 score (left panel), SNR-50 (middle panel), and SRM (right panel). Given the potential effects of age on the other variables, all subsequent analyses were completed using age as a controlling variable. Partial correlation analysis, controlling for participant age, revealed that increasing PTA was significantly correlated with poorer SRT-50, $r = .667$, $p < .0001$, and poorer SSQ-12, $r = .668$, $p < .0001$. Furthermore, poorer SSQ-12 score was significantly correlated with poorer SRT-50, $r = .671$, $p < .0001$. Conversely, SRM was not significantly correlated with any of the other unaided measures ($p > .05$). In other words, listeners with poorer hearing tended to exhibit poorer sentence recognition in noise and poorer self-assessed spatial

abilities. However, the ability to take advantage of spatial separation of speech in asymmetric competing noise was not related to these other factors.

Additional partial correlation analysis explored the possible relationship between predictive measures and bilateral benefits, as shown in Table 2. These analyses were completed while controlling statistically for both age and unilateral recognition performance since both cognition and the performance range had the potential to affect any measured relationships. Given the strong correlations between the predictive variables, we focused first on the relationships between PTA and bilateral benefits, as shown in Figures 2–5. Results revealed no significant relationships between PTA and bilateral benefits for the response times during the listening effort task (see bottom panel in Figure 4). However, PTA was significantly correlated with the benefit scores measured for AV presentation of localization speed during STRESR, $r = .43$, $p < .019$, localization accuracy during STRESR, $r = .39$, $p < .035$, word recognition with fixed talker location during the listening effort paradigm, $r = .45$, $p < .012$, and subjective spatial ratings, $r = .33$, $p < .044$. In other words, the magnitude of bilateral benefits within these AV domains generally increased with increasing hearing loss, as evidenced visually by the larger separation between unilateral and bilateral performances for higher PTAs in Figures 3–5 (right panels). Even though significant bilateral benefits were present for sentence recognition and word recognition with variable talker location during the STRESR, the magnitude of benefit was unrelated to PTA, as evidenced visually by parallel unilateral and bilateral performance lines in Figures 2 and 3.

In contrast, with AO stimuli, PTA was not significantly correlated with any of the outcomes. Not surprisingly, given the high correlations between the unaided measures of PTA, SRT-50, and SSQ-12, all three predictive measures demonstrated a similar pattern of significant and nonsignificant relationships with bilateral benefit scores with three exceptions. Specifically, the SSQ-12 score was not significantly correlated with the AV subjective spatial benefit or localization accuracy during the STRESR, but it

Figure 3. Scores for the Spatial Test Requiring Effortful Speech Recognition task with audio-only (left panel) and audiovisual (right panel) modalities. Top panels indicate word recognition performance (rau), middle panels indicate localization speed (ms), and bottom panels indicate localization accuracy (percent correct). Open symbols and dashed lines indicate performance with unilateral hearing aid; filled symbols and solid lines indicate performance with bilateral hearing aids. Data are presented as a function of better ear, pure-tone average (dB HL), calculated based on thresholds at 500, 1000, and 2000 Hz.

was significantly correlated with word recognition with variable talker location during the STRESR (see Table 2).

Discussion

The purpose of this study was to evaluate bilateral hearing aid benefit in ecologically relevant conditions, specifically across multiple auditory domains and with AO or AV stimuli. Consistent with past literature, significant bilateral benefits were identified for sentence recognition, word recognition, localization speed, localization accuracy, and subjective ratings of spatialization. Importantly, with one exception (localization accuracy), there were no interactions between bilateral hearing aid benefit and modality, demonstrating that bilateral hearing aid benefits persist even in ecologically relevant test conditions when visual cues are available. For example, an average bilateral benefit

for AV sentence recognition presented in noise of 8.7 percentage points was reported for listeners with severe hearing loss (Day et al., 1988), a value essentially identical to that found in the current study (9.13 rau). In addition, a number of studies have demonstrated significant bilateral benefits for sentences in noise using variable SNR tasks (Freyaldenhoven, Nabelek, Burchfield, & Thelin, 2005; Ricketts, 2000). The bilateral improvement of approximately 5 percentage points for word recognition in noise is also consistent with previous research (McKenzie & Rice, 1990).

Despite past studies demonstrating bilateral benefits related to subjective listening effort (Most et al., 2012; Noble, 2006; Noble & Gatehouse, 2006; van Schoonhoven et al., 2016), identification of behavioral bilateral listening effort benefits remained elusive in this study. The magnitude of differences across unilateral and bilateral reaction

Figure 4. Scores for the listening effort task with audio-only (left panel) and audiovisual (right panel) modalities. Top panels indicate word recognition performance (rau); bottom panels indicate response times during the secondary task (ms). Open symbols and dashed lines indicate performance with unilateral hearing aid; filled symbols and solid lines indicate performance with bilateral hearing aids. Data are presented as a function of better ear, pure-tone average (dB HL), calculated based on thresholds at 500, 1000, and 2000 Hz.

times was very small (< 5 ms). Thus, these results suggest that listening effort, as quantified using this dual-task paradigm, was not significantly affected by adding a second hearing aid.

The average bilateral localization advantage without visual cues of 22 percentage points was similar to, or slightly larger than, that reported previously. For example, a 10-percentage-point bilateral localization accuracy benefit has been reported for a task in which listeners identified the

location of a loudspeaker presenting a speech signal in the presence of seven other loudspeakers presenting competing noise (Köbler & Rosenhall, 2002). Conversely, a more recent study reported that root-mean-square localization error improved with the addition of the second hearing aid by approximately 23 percentage points in listeners with severe hearing loss but did not significantly improve in listeners with moderate hearing loss (van Schoonhoven et al., 2016).

Figure 5. Scores for the subjective ratings task with audio-only (left panel) and audiovisual (right panel) modalities. Higher scores indicate better self-reported spatial perception. Open symbols and dashed lines indicate performance with unilateral hearing aid; filled symbols and solid lines indicate performance with bilateral hearing aids. Data are presented as a function of better ear, pure-tone average (dB HL), calculated based on thresholds at 500, 1000, and 2000 Hz.

Figure 6. Relationship between predictive variables and pure-tone average (dB HL), calculated based on thresholds at 500, 1000, and 2000 Hz. The left panel displays the total score for the 12-item version of the Speech, Spatial and Qualities of Hearing Scale (SSQ-12); the middle panel displays the score for the estimated signal-to-noise ratio for 50%-correct recognition of key words in sentences (SNR-50) on the Quick Speech-in-Noise Test; and the right panel displays spatial release from masking (dB). The relationships between pure-tone average and the SSQ-12 and SNR-50 scores were statistically significant ($p < .0001$); the relationship between spatial release from masking and pure-tone average was not significant ($p > .05$).

The larger average benefits measured in the current study may be due to the fact that both of these previous studies implemented loudspeaker separation distances of 45° , whereas the current study included loudspeaker separations as small as 15° . Consequently, the source location identification task was generally more difficult in the current study. With AV stimuli, bilateral benefits were also evident, although the benefits were smaller (approximately 7 percentage points) than with AO stimuli, resulting in a significant interaction. These reduced localization accuracy benefits were not surprising because unilateral performance was already near ceiling for many listeners.

Specifically, 19 of the 32 listeners' localization accuracy score exceeded 90% for the unilateral condition, and only four listeners had accuracy scores lower than 70%. In other words, with visual cues present, listeners were able to locate the talker of interest eventually but took longer to do so when fitted unilaterally compared to a bilateral fitting.

Subjective bilateral spatial benefits in the current study were also similar to those reported previously. Previous research has demonstrated significant subjective bilateral advantages in localization and sound quality (Boymans et al., 2008, 2009; Köbler et al., 2001), as well as improved spaciousness (Balfour & Hawkins, 1992).

Table 2. Correlation and significance between the predictive measures and bilateral benefits for all audio-only (AO; upper panel) and audiovisual (AV; lower panel) outcomes.

		AO							
Task name	Outcome measure(s)	PTA		SSQ12		SRM		SRT-50	
		<i>r</i>	Sig.	<i>r</i>	Sig.	<i>r</i>	Sig.	<i>r</i>	Sig.
CST	Sentence recognition	-.28	.884	-.32	.076	-.17	.364	-.25	.161
STRESR	Word recognition	-.06	.114	-.28	.124	.21	.251	-.33	.169
	Localization speed	.25	.753	-.15	.407	.06	.757	-.26	.156
	Localization accuracy	.16	.894	-.39	.302	.15	.405	-.26	.312
LE	Word recognition	.28	.400	-.21	.268	.01	.942	-.25	.167
	Reaction time	.27	.271	-.14	.448	-.12	.928	-.19	.324
Subj	Subjective rating	.19	.308	-.22	.239	.12	.528	-.23	.199
		AV							
CST	Sentence recognition	-.12	.531	.07	.700	.28	.123	.02	.931
STRESR	Word recognition	.34	.075	-.42	.023	.02	.908	-.29	.113
	Localization speed	.43	.019	-.32	.022	.27	.139	-.64	.011
	Localization accuracy	.39	.035	-.30	.102	.08	.677	-.34	.041
LE	Word recognition	.45	.012	-.57	.001	.05	.782	-.65	.001
	Reaction time	.12	.942	-.22	.242	.00	.985	.10	.618
Subj	Subjective rating	.33	.044	-.27	.138	.24	.184	-.41	.021

Note. Correlations that were significant at the $p < .05$ level are presented in bold. PTA = pure-tone average; SSQ-12 = 12-item version of the Speech, Spatial and Qualities of Hearing Scale; SRM = spatial release from masking; CST = Connected Speech Test; STRESR = Spatial Test Requiring Effortful Speech Recognition; LE = Listening Effort.

Overall, these data demonstrated that, even under ecologically relevant conditions, which may be expected to reduce binaural advantages measured in the laboratory (e.g., seeing the talker's face, testing in moderate reverberation, requiring listeners to complete tasks in more than one domain in the same time, evaluating sentence recognition with considerable contextual information available), significant bilateral hearing aid advantages persisted.

Predictive Variables

A secondary purpose of this study was to determine whether the magnitude of bilateral benefits was related to participant age, degree of hearing loss, unaided sentence recognition in noise performance, unaided SRM ability, or unaided subjective spatial ability. When considering unaided measures as potential predictors of bilateral benefit, a strong association was found between poorer hearing thresholds, more difficulty recognizing speech in noise, and poorer self-reported sound quality and spatial ability as measured by the SSQ-12. These findings are in good agreement with previous studies demonstrating similar relationships (Gatehouse & Noble, 2004; Tyler, Perreau, & Ji, 2009).

Conversely, while there was a trend for decreasing SRM with increasing hearing loss, no significant relationship was found inconsistent with previous research (Marrone, Mason, & Kidd, 2008). However, the current study included a more limited range of hearing loss than represented in previous studies. Specifically, this previous study included a group of listeners with normal hearing thresholds. In addition, this same study also demonstrated that SRM decreases with increasing age (Marrone et al., 2008). Therefore, the lack of a significant relationship between SRM and PTA in the current study may be attributable to a more homogeneous sample with regard to age and hearing loss. The lack of correlation may also have been attributable to the use of asymmetric noise in the current study, whereas Marrone et al. (2008) used a symmetric noise configuration. The HINT was chosen in the current study because it is clinically viable (requiring only two speakers or headphones) and has considerable normative data for both collocated and separated speech-in-noise conditions (S. D. Soli & Wong, 2008). However, the asymmetric noise configuration provides a head-shadow (SNR) advantage in addition to true binaural speech-in-noise advantages. Consequently, the head-shadow advantage likely dominated the SRM results, limiting the relationship between this measure and bilateral benefits.

Given the strong relationship between the other three of the predictive measures, it was not surprising that all of them (PTA, SSQ-12, and SRT-50) were related in a similar way to bilateral benefits. Given this similarity and to simplify comparisons, we focus mainly on the relationships between calculated bilateral benefits and degree of hearing loss (PTA). When controlling for age and unilateral speech recognition performance, the relationship between bilateral benefits and hearing loss is different with AO and AV stimuli (see Table 2).

With AO stimuli, none of the outcomes was related to PTA. The limited relationship between PTA and bilateral benefits with AO stimuli appears to be in contrast to previous studies demonstrating significant relationships (e.g., van Schoonhoven et al., 2016). However, a closer examination of previous studies reveals that participants were typically grouped into those with milder and more severe hearing loss. Furthermore, the milder group typically was categorized as having PTAs of better than 40 dB HL (van Schoonhoven et al., 2016). Although the current study had significant variability in PTA, potential hearing aid users were targeted; thus, hearing thresholds typically fell in the moderate-to-severe range. Hearing aid use in listeners with mild degrees of hearing loss is known to be very low (Chien & Lin, 2012; Ciletti & Flamme, 2008). Only seven participants had PTA thresholds better than 40 dB HL in the current study. In good agreement with previous research, the bilateral benefits in this small subset of participants with mild hearing loss were generally negligible, except for the sentence recognition task (see Figure 2). Consequently, our results would likely be very similar to this previous report if participants were categorized into “mild” versus “moderate and greater” hearing loss groups rather than performing correlation analysis.

The current AO findings suggest that, for listeners with PTAs greater than 40 dB HL, the magnitude of the measured bilateral benefits is not significantly related to degree of hearing loss. That is, while there is considerable individual variability, the magnitude of average AO bilateral benefit is negligible when hearing is near normal and then increases. The left panels in Figures 3–5 show increasing separation between the trend lines on several outcomes. From these figures, it appears that bilateral benefits increase with increasing hearing loss. However, this increase does not reach significance after controlling for age and unilateral recognition performance.

For AV stimuli, the data demonstrated a weak, but significant, relationship between degree of hearing loss and bilateral benefits on several outcomes (see Figures 3–5 and Table 2). The significant relationship between increasing PTA and increasing AV bilateral benefits was present even when controlling for age differences, suggesting that the differences are due to accessibility of sound at both ears, rather than age-related cognitive factors. One potential consideration is the performance range over which bilateral benefits were evaluated. As is evident from Figures 2–5, listeners with less hearing loss were evaluated in a performance range that was higher than that for those with greater degrees of hearing loss. This is a potential confound because visual speech recognition benefits are typically greatest at very poor SNRs in speech with limited context (e.g., Grant & Seitz, 1998). However, the significant relationship between increasing PTA and increasing AV bilateral benefits continued even when controlling for unilateral performance, suggesting that it was not due to the established increase in visual benefits when performance is poorest. In addition, the individuals with the greatest hearing loss and poorest performance were also evaluated at the most positive SNRs

(where visual benefits would be expected to be the smallest). Based on the current study design, performance range cannot be completely ruled out as a possible contributing factor. For example, the relationship between word recognition and PTA was not significant when the talker location was variable (during the STRESR) but was significant when fixed (during listening effort testing). Whereas these two outcomes were measured at the same SNRs within an individual, the average performance range was very different (see top panel in Figure 3 vs. top panel in Figure 4). Therefore, the lack of relationship between PTA and AV word recognition with variable location could be due to the lower performance range for the STRESR. Specifically, listeners in the poorest performance range (highest PTA while completing the STRESR task) would be expected to obtain the largest visual benefits. Consequently, they might be expected to decreasingly rely on auditory information, potentially decreasing the potential for bilateral benefits with increasing hearing loss. However, there are other possible explanations for the lack of significant relationship between PTA and AV word recognition collected as part of the STRESR. Specifically, unlike simple word recognition, the STRESR requires word recall and includes a variable source location. It may be that the complexity of this task reduces the benefits of adding the second hearing aid on the word recognition component.

One explanation for the significant relationship between bilateral benefits and hearing loss for AV presentations, but not AO, may relate to factors affecting sensory integration (Bernstein & Grant, 2009) and SNR differences between the conditions. Visual cues improve speech intelligibility in noise, both by providing additional phonetic information such as place of articulation (Auer & Bernstein, 2008) and by improving segregation of speech and noise sources (Wightman, Kistler, & Brungart, 2006). Since the addition of a second hearing aid does not greatly improve audibility, the bilateral benefits measured in the current study might be primarily due to improved source segregation. However, some benefit could also occur because bilateral summation may increase perceived sound levels.

Source segregation can be enhanced by the co-modulation of visual and acoustic segments of speech and is largest for fluctuating maskers (Bernstein & Grant, 2009), such as the four-talker babble used in the current study. In the case of AV bilateral benefits, source segregation may be further enhanced through improved spatial segregation of speech and noise sources resulting from improved localization. Therefore, it is possible that, with increasing PTA, audibility of spatial cues for the unaided ear decreased for both AO and AV conditions. This, in turn, results in greater increases in access to spatial cues with the addition of the second hearing aid as PTA increases. However, in the AV conditions, a poorer SNR was used in comparison to the AO conditions to maintain similar performance. This lower SNR was expected to further degrade access to the audio information in the unaided ear of the unilateral conditions and, therefore, increase benefits when adding the second hearing aid. This degradation in

unaided audio information may also have increased the importance of accurate localization for improving spatial source segregation. Further work in this area is needed to test this hypothesis.

It is also unclear why AV sentence recognition benefit was not significantly related to PTA. One possibility for this lack of correlation may be related to the additional contextual cues provided when using the topic-based sentence materials of the CST. It is well established that increasing linguistic context improves speech recognition performance in noise and other adverse conditions (Nittrouer & Boothroyd, 1990; Pichora-Fuller, 2008). However, data have also shown that the benefits associated with increasing context decrease with increasing hearing loss (Meister, Schreitmüller, Ortmann, Rähmann, & Walger, 2016). Therefore, the increases in bilateral speech recognition benefits with increasing PTA could be offset by decreases in contextual benefits.

Another possibility relates to the average benefits measured for those with hearing thresholds better than 40 dB HL, who demonstrate negligible bilateral benefits measured across most domains (see Figures 3–5). In contrast, bilateral benefits for sentence recognition in noise in these same listeners were approximately equivalent to the average measured for all participants (approximately 7–8 percentage points). It may be that the lower presentation level used for the CST in listeners with better hearing thresholds may have increased the benefit gained from the second hearing aid even in listeners with milder hearing loss.

Conclusion

In conclusion, the results of this study demonstrate that, when SNR is adjusted to provide equivalent performance with AO and AV stimuli for unilateral hearing aid conditions, average bilateral benefits are present and generally independent of modality. This finding was consistent across the domains of speech recognition, localization speed, and subjective spatial qualities. While considerable individual variability is present, hearing aid users can therefore expect significant bilateral hearing aid advantages even in ecologically relevant conditions. Furthermore, when considering degree of hearing loss consistent with listeners who pursue amplification (PTAs > 40 dB HL), the magnitude of average bilateral benefits was similar across a wide range of PTAs for AV sentence recognition and all AO domains (word recognition, localization speed, localization accuracy, subjective spatial ratings). In contrast, the data demonstrate a significant relationship between degree of hearing loss and bilateral benefits with AV stimuli across many domains (localization speed, localization accuracy, subjective spatial ratings); listeners with more hearing loss benefited more from the second hearing aid. Overall, these data suggest the presence of vision strengthens the relationship between bilateral benefits and degree of hearing loss, although further work is necessary to determine if these results may be attributable to differences in SNR for the AO versus AV modalities. Finally, the data do not support smaller bilateral

benefits in ecologically relevant test conditions as a possible explanation for the discrepancy between noted laboratory bilateral benefits and patient preferences for two hearing aids.

Acknowledgments

This research was funded by a grant from GN ReSound (PI: Todd Ricketts), by the Maddox Charitable Trust, and by National Institute on Deafness and Other Communication Disorders Grant T35 DC008763 (awarded to the third author; PI: Linda Hood). The authors wish to thank Susan Stangl, Elizabeth Harland, Jodi Rokuson, and Kristen D'Onofrio for their efforts in recruiting and collecting data. Portions of this project were presented at the Scientific and Technical Meeting of the American Auditory Society (March 2013).

References

- Allen, R. L., Schwab, B. M., Cranford, J. L., & Carpenter, M. D. (2000). Investigation of binaural interference in normal-hearing and hearing-impaired adults. *Journal of the American Academy of Audiology*, 11(9), 494–500.
- Auer, E. T., Jr., & Bernstein, L. E. (2008). Estimating when and how words are acquired: A natural experiment on the development of the mental lexicon. *Journal of Speech, Language, and Hearing Research*, 51(3), 750–758.
- Balfour, P. B., & Hawkins, D. B. (1992). A comparison of sound quality judgments for monaural and binaural hearing aid processed stimuli. *Ear and Hearing*, 13(5), 331–339.
- Bernstein, J. G., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 125(5), 3358–3372.
- Bertoli, S., Bodmer, D., & Probst, R. (2010). Survey on hearing aid outcome in Switzerland: Associations with type of fitting (bilateral/unilateral), level of hearing aid signal processing, and hearing loss. *International Journal of Audiology*, 49(5), 333–346.
- Best, V., Mason, C. R., & Kidd, G., Jr. (2011). Spatial release from masking in normally hearing and hearing-impaired listeners as a function of the temporal overlap of competing talkers. *The Journal of the Acoustical Society of America*, 129(3), 1616–1625.
- Boymans, M., Goverts, S. T., Kramer, S. E., Festen, J. M., & Dreschler, W. A. (2008). A prospective multi-centre study of the benefits of bilateral hearing aids. *Ear and Hearing*, 29(6), 930–941.
- Boymans, M., Goverts, S. T., Kramer, S. E., Festen, J. M., & Dreschler, W. A. (2009). Candidacy for bilateral hearing aids: A retrospective multicenter study. *Journal of Speech, Language, and Hearing Research*, 52(1), 130–140.
- Brody, L., Wu, Y. H., & Stangl, E. (2018). A comparison of personal sound amplification products and hearing aids in ecologically relevant test environments. *American Journal of Audiology*, 27(4), 581–593.
- Byrne, D., Dillon, H., Ching, T., Katsch, R., & Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. *Journal of the American Academy of Audiology*, 12(1), 37–51.
- Byrne, D., Noble, W., & LePage, B. (1992). Effects of long-term bilateral and unilateral fitting of different hearing aid types on the ability to locate sounds. *Journal of the American Academy of Audiology*, 3(6), 369–382.
- Carter, A. S., Noe, C. M., & Wilson, R. H. (2001). Listeners who prefer monaural to binaural hearing aids. *Journal of the American Academy of Audiology*, 12(5), 261–272.
- Chien, W., & Lin, F. R. (2012). Prevalence of hearing aid use among older adults in the United States. *Archives of Internal Medicine*, 172(3), 292–293.
- Chmiel, R., Jerger, J., Murphy, E., Pirozzolo, F., & Tooley-Young, C. (1997). Unsuccessful use of binaural amplification by an elderly person. *Journal of the American Academy of Audiology*, 8(1), 1–10.
- Chung, S., & Stephens, S. (1986). Factors influencing binaural hearing aid use. *British Journal of Audiology*, 20(2), 129–140.
- Ciletti, L., & Flamme, G. A. (2008). Prevalence of hearing impairment by gender and audiometric configuration: Results from the National Health and Nutrition Examination Survey (1999–2004) and the Keokuk County Rural Health Study (1994–1998). *Journal of the American Academy of Audiology*, 19(9), 672–685.
- Cox, R. M., Alexander, G., & Gilmore, C. (1987a). Development of the Connected Speech Test (CST). *Ear and Hearing*, 8(Suppl. 5), 119S–126S.
- Cox, R. M., Alexander, G. C., & Gilmore, C. (1987b). Intelligibility of average talkers in typical listening environments. *The Journal of the Acoustical Society of America*, 81(5), 1598–1608.
- Cox, R. M., Alexander, G. C., Gilmore, C., & Pusakulich, K. M. (1989). The Connected Speech Test version 3: Audiovisual administration. *Ear and Hearing*, 10(1), 29–32.
- Cox, R. M., Schwartz, K. S., Noe, C. M., & Alexander, G. C. (2011). Preference for one or two hearing aids among adult patients. *Ear and Hearing*, 32(2), 181–197.
- Day, G. A., Browning, G. G., & Gatehouse, S. (1988). Benefit from binaural hearing aids in individuals with a severe hearing impairment. *British Journal of Audiology*, 22(4), 273–277.
- Dillon, H., Birtles, G., & Lovegrove, R. (1999). Measuring the outcomes of a national rehabilitation program: Normative data for the client oriented scale of improvement (COSI) and the Hearing Aid User's Questionnaire (HAUQ). *Journal of the American Academy of Audiology*, 10(2), 67–79.
- Downs, D. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders*, 47(2), 189–193.
- Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293), 52–64.
- Durlach, N. I., Thompson, C. L., & Colburn, H. S. (1981). Binaural interaction in impaired listeners: A review of past research. *International Journal of Audiology*, 20(3), 181–211.
- Erber, N. P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders*, 40(4), 481–492.
- Erdman, S. A., & Sedge, R. K. (1981). Subjective comparisons of binaural versus monaural amplification. *Ear and Hearing*, 2(5), 225–229.
- Festen, J. M., & Plomp, R. (1986). Speech-reception threshold in noise with one and two hearing aids. *The Journal of the Acoustical Society of America*, 79(2), 465–471.
- Feuerstein, J. (1992). Monaural versus binaural hearing: Ease of listening, word recognition, and attentional effort. *Ear and Hearing*, 13(2), 80–86.
- Fraser, S., Gagné, J., Alepins, M., & Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. *Journal of Speech, Language, and Hearing Research*, 53(1), 18–33.
- Freyaldenhoven, M. C., Nabelek, A. K., Burchfield, S. B., & Thelin, J. W. (2005). Acceptable noise level as a measure of directional hearing aid benefit. *Journal of the American Academy of Audiology*, 16(4), 228–236.

- Freyaldenhoven, M. C., Plyler, P. N., Thelin, J. W., & Burchfield, S. B. (2006). Acceptance of noise with monaural and binaural amplification. *Journal of the American Academy of Audiology*, 17(9), 659–666.
- Gatehouse, S., Elberling, C., & Naylor, G. (1999). Aspects of auditory ecology and psychoacoustic function as determinants of benefits from and candidature for non-linear processing in hearing aids. In A. N. Rasmussen & P. A. Osterhammel (Eds.), *Auditory models and non-linear hearing instruments* (pp. 221–233). Kolding, Denmark: Danavox Jubilee Foundation.
- Gatehouse, S., & Noble, W. (2004). The Speech, Spatial and Qualities of Hearing Scale (SSQ). *International Journal of Audiology*, 43(2), 85–99.
- Grant, K. W., & Seitz, P. F. (1998). Measures of auditory–visual integration in nonsense syllables and sentences. *The Journal of the Acoustical Society of America*, 104(4), 2438–2450.
- Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory–visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America*, 103(5), 2677–2690.
- Guyatt, G. H., Oxman, A. D., Kunz, R., Brozek, J., Alonso-Coello, P., Rind, D., ... Schünemann, H. J. (2011). GRADE guidelines 6. Rating the quality of evidence—Imprecision. *Journal of Clinical Epidemiology*, 64(12), 1283–1293.
- Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., ... Vist, G. (2011). GRADE guidelines: 8. Rating the quality of evidence—Indirectness. *Journal of Clinical Epidemiology*, 64(12), 1303–1310.
- Guyatt, G. H., Oxman, A. D., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., ... Falck-Ytter, Y. (2011). GRADE guidelines: 4. Rating the quality of evidence—Study limitations (risk of bias). *Journal of Clinical Epidemiology*, 64(4), 407–415.
- Hall, E. T. (1966). *The hidden dimension*. New York, NY: Doubleday.
- Hawkins, D. B., & Yacullo, W. S. (1984). Signal-to-noise ratio advantage of binaural hearing aids and directional microphones under different levels of reverberation. *Journal of Speech and Hearing Disorders*, 49(3), 278–286.
- Hedgecock, L. D., & Sheets, B. V. (1958). A comparison of monaural and binaural hearing aids for listening to speech. *AMA Archives of Otolaryngology*, 68(5), 624–629.
- Helfer, K. S., & Freyman, R. L. (2005). The role of visual speech cues in reducing energetic and informational masking. *The Journal of the Acoustical Society of America*, 117(2), 842–849.
- Hicks, C., & Tharpe, A. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*, 45(3), 573–584.
- Jackson, C. (1953). Visual factors in auditory localization. *The Quarterly Journal of Experimental Psychology*, 5(2), 52–65.
- Jerger, J., Carhart, R., & Dirks, D. (1961). Binaural hearing aids and speech intelligibility. *Journal of Speech and Hearing Research*, 4(2), 137–148.
- Jerger, J., Silman, S., Lew, H. L., & Chmiel, R. (1993). Case studies in binaural interference: Converging evidence from behavioral and electrophysiologic measures. *Journal of the American Academy of Audiology*, 4(2), 122–131.
- Killion, M. C., Niquette, P. A., Revit, L. J., & Skinner, M. W. (2001). Quick SIN and BKB-SIN, two new speech-in-noise tests permitting SNR-50 estimates in 1 to 2 min. *The Journal of the Acoustical Society of America*, 109, 2502.
- Köbler, S., Lindblad, A. C., Olofsson, Å., & Hagerman, B. (2010). Successful and unsuccessful users of bilateral amplification: Differences and similarities in binaural performance. *International Journal of Audiology*, 49(9), 613–627.
- Köbler, S., & Rosenhall, U. (2002). Horizontal localization and speech intelligibility with bilateral and unilateral hearing aid amplification: Localización horizontal y discriminación del lenguaje con adaptación unilateral y bilateral de auxiliares auditivos. *International Journal of Audiology*, 41(7), 395–400.
- Köbler, S., Rosenhall, U., & Hansson, H. (2001). Bilateral hearing aids—Effects and consequences from a user perspective. *Scandinavian Audiology*, 30(4), 223–235.
- Markle, D. M., & Aber, W. (1958). A clinical evaluation of monaural and binaural hearing aids. *AMA Archives of Otolaryngology*, 67(5), 606–608.
- Marrone, N., Mason, C. R., & Kidd, G., Jr. (2008). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *The Journal of the Acoustical Society of America*, 124(5), 3064–3075.
- McArdle, R. A., Killion, M., Mennite, M. A., & Chisolm, T. H. (2012). Are two ears not better than one. *Journal of the American Academy of Audiology*, 23(3), 171–181.
- McKenzie, A. R., & Rice, C. (1990). Binaural hearing aids for high-frequency hearing loss. *British Journal of Audiology*, 24(5), 329–334.
- Meister, H., Schreitmüller, S., Ortmann, M., Rähmann, S., & Walger, M. (2016). Effects of hearing loss and cognitive load on speech recognition with competing talkers. *Frontiers in Psychology*, 7, 301.
- Miller, C. W., Stewart, E. K., Wu, Y. H., Bishop, C., Bentler, R. A., & Tremblay, K. (2017). Working memory and speech recognition in noise under ecologically relevant listening conditions: Effects of visual cues and noise type among adults with hearing loss. *Journal of Speech, Language, and Hearing Research*, 60(8), 2310–2320.
- Moore, T. M., & Picou, E. M. (2018). A potential bias in subjective ratings of mental effort. *Journal of Speech, Language, and Hearing Research*, 61, 2405–2421.
- Most, T., Adi-Bensaid, L., Shpak, T., Sharkiya, S., & Luntz, M. (2012). Everyday hearing functioning in unilateral versus bilateral hearing aid users. *American Journal of Otolaryngology*, 33(2), 205–211.
- Mussoi, B. S., & Bentler, R. A. (2017). Binaural interference and the effects of age and hearing loss. *Journal of the American Academy of Audiology*, 28(1), 5–13.
- Naidoo, S. V., & Hawkins, D. B. (1997). Monaural/binaural preferences: Effect of hearing aid circuit on speech intelligibility and sound quality. *Journal of the American Academy of Audiology*, 8(3), 188–202.
- Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America*, 87(6), 2705–2715.
- Noble, W. (2006). Bilateral hearing aids: A review of self-reports of benefit in comparison with unilateral fitting: Auxiliares auditivos bilaterales: Revisión de los beneficios auto-reportados en comparación con la adaptación unilateral. *International Journal of Audiology*, 45(Suppl. 1), 63–71.
- Noble, W., & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the Speech, Spatial, and Qualities of Hearing scale (SSQ). *International Journal of Audiology*, 45(3), 172–181.
- Noble, W., Jensen, N. S., Naylor, G., Bhullar, N., & Akeroyd, M. A. (2013). A short form of the Speech, Spatial and Qualities of Hearing scale suitable for clinical use: The SSQ12. *International Journal of Audiology*, 52(6), 409–412.

- O'Neill, J. J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech and Hearing Disorders*, 19(4), 429–439.
- Pearsons, K. S., Bennett, R., & Fidell, S. (1977). *Speech levels in various noise environments* (Vol. EPA-600/1-77-025). Washington, DC: U.S. Environmental Protection Agency.
- Peutz, V. (1971). Articulation loss of consonants as a criterion for speech transmission in a room. *Journal of the Audio Engineering Society*, 19(11), 915–919.
- Pichora-Fuller, K. M. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, 47(Suppl. 2), S72–S82.
- Picou, E. M., Aspell, E., & Ricketts, T. A. (2014). Potential benefits and limitations of three types of directional processing in hearing aids. *Ear and Hearing*, 35(3), 339–352.
- Picou, E. M., Ricketts, T. A., & Hornsby, B. W. (2011). Visual cues and listening effort: Individual variability. *Journal of Speech, Language, and Hearing Research*, 54, 1416–1430.
- Picou, E. M., Ricketts, T. A., & Hornsby, B. W. (2013). How hearing aids, background noise, and visual cues influence objective listening effort. *Ear and Hearing*, 34, e52–e64.
- Punch, J. L., Jenison, R. L., Allan, J., & Durrant, J. D. (1991). Evaluation of three strategies for fitting hearing aids binaurally. *Ear and Hearing*, 12(3), 205–215.
- Ricketts, T. (2000). The impact of head angle on monaural and binaural performance with directional and omnidirectional hearing aids. *Ear and Hearing*, 21(4), 318–328.
- Schreurs, K. K., & Olsen, W. O. (1985). Comparison of monaural and binaural hearing aid use on a trial period basis. *Ear and Hearing*, 6(4), 198–202.
- Silberer, A. B., Bentler, R., & Wu, Y. H. (2015). The importance of high-frequency audibility with and without visual cues on speech recognition for listeners with normal hearing. *International Journal of Audiology*, 54, 865–872.
- Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology*, 26(2), 183–196.
- Soli, S., & Nilsson, M. (1994). Assessment of communication handicap with the HINT. *Hearing Instruments*, 45, 12.
- Soli, S. D., & Wong, L. L. (2008). Assessment of speech intelligibility in noise with the Hearing in Noise Test. *International Journal of Audiology*, 47(6), 356–361.
- Stephens, S. D., Callaghan, D. E., Hogan, S., Meredith, R., Rayment, A., & Davis, A. (1991). Acceptability of binaural hearing aids: A cross-over study. *Journal of the Royal Society of Medicine*, 84(5), 267–269.
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28(3), 455–462.
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26(2), 212–215.
- Tyler, R. S., Perreau, A. E., & Ji, H. (2009). Validation of the Spatial Hearing Questionnaire. *Ear and Hearing*, 30(4), 466–474.
- van Schoonhoven, J., Schulte, M., Boymans, M., Wagener, K. C., Dreschler, W. A., & Kollmeier, B. (2016). Selecting appropriate tests to assess the benefits of bilateral amplification with hearing aids. *Trends in Hearing*, 20. <https://doi.org/10.1177/2331216516658239>
- Vaughan-Jones, R. H., Padgham, N. D., Christmas, H. E., Irwin, J., & Doig, M. (1993). One aid or two?—More visits please! *The Journal of Laryngology & Otology*, 107(04), 329–332.
- Walden, B. E. (1997). Toward a model clinical-trials protocol for substantiating hearing aid user-benefit claims. *American Journal of Audiology*, 6(2), 13–24.
- Walden, B. E., & Walden, T. C. (2005). Unilateral versus bilateral amplification for adults with impaired hearing. *Journal of the American Academy of Audiology*, 16(8), 574–584.
- Walden, T. C., & Walden, B. E. (2004). Predicting success with hearing aids in everyday living. *Journal of the American Academy of Audiology*, 15(5), 342–352.
- Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88(3), 638–687.
- Wightman, F., Kistler, D., & Brungart, D. (2006). Informational masking of speech in children: Auditory–visual integration. *The Journal of the Acoustical Society of America*, 119, 3940–3949.
- Wu, Y. H., & Bentler, R. A. (2010). Impact of visual cues on directional benefit and preference: Part I—Laboratory tests. *Ear and Hearing*, 31(1), 22–34.
- Wu, Y. H., Stangl, E., Chipara, O., Hasan, S. S., Welhaven, A., & Oleson, J. (2018). Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss. *Ear and Hearing*, 39(2), 293–304.