Original Article

Trends in Hearing
2016, Vol. 20: 1–16

DOI: 10.1177/2331216516658239
tia.sagepub.com
SAGE

Selecting Appropriate Tests to Assess the Benefits of Bilateral Amplification With Hearing Aids

Jelmer van Schoonhoven[1], Michael Schulte[2,3], Monique Boymans[1], Kirsten C. Wagener[2,3], Wouter A. Dreschler[1], and Birger Kollmeier[2,3,4]

Abstract

The aim of this study was to investigate the effect of bilateral hearing aids (HA) in subjects with mild and moderate-to-severe hearing loss. This study was designed as a within-subject feasibility study. Bilateral HA use was assessed using different laboratory tests on speech reception, listening effort, noise tolerance, and localization. All data were evaluated with bilateral and unilateral HA fittings. Forty experienced bilateral HA users were included with hearing impairment ranging from mild to moderate-to-severe. Subjects were stratified into two groups based on the degree of hearing loss. Speech reception in noise, listening effort, and localization tests showed a bilateral benefit for the moderate-to-severely hearing-impaired subjects. A bilateral benefit was also observed for listening effort in the mildly hearing-impaired group. The assessment of listening effort shows promise as a measure of bilateral HA benefit for mild hearing impairment. Localization and speech reception in noise tests provide additional value for larger losses. The next step is to compare experienced unilateral with bilateral HA users.

Keywords

hearing loss, hearing aids, bilateral hearing aids, speech intelligibility in noise, listening effort, sound localization, sound detection, psychoacoustics

Date received: 24 February 2016; revised: 14 June 2016; accepted: 14 June 2016

Introduction

Hearing with two normal ears has several advantages over monaural hearing. These advantages include better speech reception in noise, especially when speech and noise are spatially separated (Persson, Harder, Arlinger, & Magnuson, 2001; Plomp, 1976), a reduction in listening effort in certain noise conditions (Feuerstein, 1992), and better horizontal localization (Grothe, Pecka, & McAlpine, 2010; Irving & Moore, 2011; Middlebrooks & Green, 1991). However, binaural advantages for subjects fitted with bilateral hearing aids (HAs) are less clear (e.g., Freyaldenhoven, Plyler, Thelin, & Burchfield, 2006; Kim, Lee, & Lee, 2014; McArdle, Killion, Mennite, & Chisolm, 2012; Walden & Walden, 2005). Results between studies about the benefit of bilateral versus unilateral amplification using HAs do not always align and are at times contradictory. Furthermore, data on the effect of bilateral HAs in the domain of listening effort are scarce. Finally, little is reported about tests that aim to find bilateral benefit in real life or simulated real life conditions, as opposed to the traditional laboratory tests. The goal of the current study was, therefore, to assess the added value of a second HA on different dimensions of performance: speech reception in noise, listening effort, noise tolerance, and localization. In the text later, the issue of bilateral amplification is addressed for each of these four domains.

Speech Reception in Noise

When speech and noise are identical at both ears (diotic stimulation), it is expected that any binaural benefit is the result of binaural redundancy. For normally hearing

[1]Department of Clinical and Experimental Audiology, Academic Medical Centre, Amsterdam, The Netherlands
[2]Hörzentrum Oldenburg GmbH, Oldenburg, Germany
[3]Cluster of Excellence Hearing4all, Oldenburg, Germany
[4]Medizinische Physik, Carl-von-Ossietzky Universität Oldenburg, Germany

Corresponding Author:
Jelmer van Schoonhoven, D2-226, Clinical and Experimental Audiology, Academic Medical Centre, Meibergdreef 9, 1105AZ Amsterdam, The Netherlands.
Email: jvanschoonhoven@amc.nl

subjects, this leads to an improvement in the speech reception threshold (SRT) of 1 to 2 dB (Moore, Johnson, Clark, & Pluvinage, 1992; Plomp, 1976). Walden and Walden (2005) and Henkin, Waldman, and Kishon-Rabin (2007) presented speech from the front and noise from the back at 180° and found a disadvantage of a second HA in the majority of elderly hearing-impaired subjects (82% and 71% in the two studies, respectively). In contrast, McArdle et al. (2012) repeated the study of Walden and Walden with subjects of a similar age and found that 80% of the subjects performed better with bilateral versus unilateral amplification. The slightly younger subjects tested by Freyaldenhoven et al. (2006) with a similar setup showed an average a bilateral benefit of 3.3 dB.

By spatially separating the sources (e.g., placing the noise source under an angle, resulting in dichotic stimulation), both binaural squelch and the head shadow effect can play a role. The head functions as a sound baffle, creating an acoustic shadow. When a sound source is placed under an angle, a listener profits from the head shadow effect by attending to the ear with the better (signal to noise ratio) SNR. Binaural squelch is the result of centrally combining the signals presented at both ears. In the case of dichotic stimulation, this may lead to better performance when listening with two ears instead of one. According to Bronkhorst and Plomp (1989), the head shadow effect results in an increase in intelligibility up to 8 dB, whereas binaural squelch leads to an improvement of around 5 dB. These values were found in normally hearing subjects.

Using different dichotic configurations, a bilateral HA benefit between 3 and 7 dB has been reported in the literature (Boymans, Goverts, Kramer, Festen, & Dreschler, 2008; Festen & Plomp, 1986; Kobler & Rosenhall, 2002; Markides, 1982). Festen and Plomp (1986) further mentioned that the head shadow effect does not apply for mild hearing loss in combination with high noise levels, since speech in the unaided ear is sufficiently audible. Other factors besides hearing loss and loudspeaker configuration that may influence the benefit of a second HA are signal characteristics, reverberation, position of the transducers, and HA configuration.

Noble (2006) conducted a review of 14 studies concerning self-reports about the benefit of bilateral HA fittings. Bilateral HAs were found to offer no advantage in situations with relatively stationary competing noise. However, a benefit was reported on the Speech, Spatial and Qualities of Hearing Scale (SSQ) in situations with switching speech streams, rapidly switching and divided attention, and listening effort (Noble & Gatehouse, 2006). Similar benefits on the SSQ were found by Most, Adi-Bensaid, Shpak, Sharkiya, and Luntz (2012).

Due to the contradictory results in the literature on speech reception in noise using diotic stimulation, we chose to include diotic stimuli in our test battery. In addition, we evaluated speech reception with speech from the front and stationary speech shaped noise from either the unilaterally aided side or from the unilaterally unaided side. These dichotic configurations made it possible to evaluate whether the head shadow effect or binaural squelch played a role. Besides this classic test setup, an interleaved speech reception test with switching speech and noise sources was used in the current study. This setup was chosen based on the self-reported findings of Noble and Gatehouse (2006), where a bilateral HA benefit was found to be most pronounced in dynamic listening situations.

Listening Effort

Little has been reported about the effect of bilateral HA fittings on listening effort. Feuerstein (1992) tested the performance of normally hearing subjects on speech reception in noise using dichotic stimulation in monaural and binaural conditions. At the same time, he assessed both ease of listening (using a 100-point scale) and attentional effort (using a dual task paradigm). He concluded that a mild simulated conductive hearing loss reduced the subjectively rated ease of listening, even when the noise source was on the side of the non-attenuated ear. Noble and Gatehouse (2006) used the SSQ questionnaire on unilateral and bilateral HA users and found a significant reduction in “effort required to engage in the activity of listening in the everyday world” when adding a second HA (a reduction of 1.53 on a scale from 0 to 10). Most et al. (2012) reported similar findings. To investigate the effect of a second HA on listening effort, Listening Effort Scaling (LES) was included in the current study. In this test, subjects are asked to indicate the amount of effort it takes to listen to a sound (generally speech). The scale ranges from 0 (*no effort*) to 6 (*extreme effort*).

Acceptable Noise Level

The Acceptable Noise Level (ANL) is a test to investigate what level of noise is tolerated while listening to continuous discourse. Wu, Stangl, Pang, and Zhang (2014) found a 1.9 dB binaural benefit in normally hearing subjects using diotic stimulation, but no benefit using dichotic stimulation. This finding was unexpected, since dichotic stimulation resulted in better speech reception in noise. Freyaldenhoven et al. (2006) compared speech reception in noise and the ANL with one and two HAs when presenting speech from the front and noise from the back (180°). An improvement in SRT was found, but a second HA did not affect the acceptance of noise. On the other hand, Kim et al. (2014) used one frontal loudspeaker for both speech and noise and found a small but significant bilateral benefit of 1.6 dB in ANL, averaged across different types of noise. Based on the available literature, no clear conclusions can be drawn about the impact of binaural hearing or bilateral amplification on the ANL. In the current study, we, therefore, chose to include this outcome measure.

Localization and Spatial Detection

Irving and Moore (2011) conducted localization experiments in normally hearing subjects with earplugs to simulate a mild to moderate unilateral conductive hearing loss. Although localization abilities in the monaural condition improved as a result of training, performance remained significantly poorer compared with the binaural condition. Kobler and Rosenhall (2002) tested 19 experienced bilateral HA users with a mild-to-moderate sensorineural hearing loss and found that bilateral amplification improved localization compared with unilateral amplification. Similarly, Byrne, Noble, and LePage (1992) found that the addition of a second HA improved localization, as did Punch, Jenison, Allan, and Durrant (1991). Byrne et al. further stated that bilateral amplification only improved localization abilities in subjects with moderate or severe hearing loss and not in those with mild hearing loss. The self-reported results published by Noble, Ter-Horst, and Byrne (1995) point in the same direction, since they found no advantage of a second HA in subjects with a mild hearing loss. Noble and Gatehouse (2006) reported that bilateral HAs only show benefit over a unilateral HA in dynamic areas of spatial hearing, such as movement discrimination. Vaughan-Jones, Padgham, Christmas, Irwin, and Doig (1993) made use of self-reports and found that a second HA was disadvantageous for localization. Using a crossover design testing unilateral and bilateral amplification, they provided subjects with a HA questionnaire during multiple visits. Finally, Akeroyd (2014) sums up various studies that all reported larger localization errors with two HAs than without HAs, even after 3 to 15 weeks of acclimatization. In the current study, localization abilities were assessed in a complex sound field in order to mimic more dynamic daily life situations. This setup was chosen based on the aforementioned self-reported findings by Noble and Gatehouse (2006) and to investigate whether this benefit was also seen in the laboratory. Besides localization, the same setup was also used for a spatial detection task.

Rationale for Present Study

A test battery was designed that focused on the previously mentioned domains and was implemented to investigate whether a bilateral benefit could be demonstrated. This study was designed as a within-subject feasibility study using experienced bilateral HA users. The aim of the current study was to investigate which dimensions of performance show benefit from bilateral HAs. NH subjects were also tested in order to get an indication of the maximum possible bilateral benefit.

An important issue is whether this study should be conducted with HA users who are used to wearing two HAs. Trials by Erdman and Sedge (1981), Schreurs and Olsen (1985), Day, Browning, and Gatehouse (1988), Vaughan-Jones et al. (1993), and Cox, Schwartz, Noe, and Alexander (2011) showed that between 20% and 61% of the subjects who compared unilateral and bilateral HA fittings eventually chose to wear one HA. Furthermore, Noble and Byrne (1991) stated that their differences in outcomes are best accounted for by patterns of HA use, rather than by test conditions.

The goals of the current study were to find out whether the binaural benefit is retained in HA users and which tests are the most sensitive to quantify the effects. Our choice to conduct the experiments with experienced bilateral HA users inevitably means that subjects who prefer unilateral amplification were not included in this study. It is important to note that this choice could have introduced a bias toward bilateral benefit. This bias will be addressed further in the Discussion section.

Materials and Methods

An (international) two-center study protocol was used, in part to ensure that the results were not determined by a specific test setup in one center. The study was conducted at the Academic Medical Centre, Amsterdam, the Netherlands (AMC) and at the Hörzentrum Oldenburg, Germany (HZO). Given the different languages in the two centers, outcomes that are found systematically across centers are likely to be generally applicable. This would an important result for international multicenter projects, especially in Europe with its many languages.

Subjects

Forty subjects with sensorineural hearing loss were included, all of whom had more than 1 year of experience with bilateral HAs and used them for more than 5 hours per day (based on self-report). The subjects were evenly distributed over the two centers and had a mean age of 55 years (range: 23–68 years) in the AMC and 70 years (range: 54–84 years) in the HZO.

The hearing thresholds were symmetrical. Symmetry was defined as a left-right difference of $\leqslant 10\,\mathrm{dB}$ in pure-tone average ($PTA_{(0.5,1,2,4\,\mathrm{kHz})}$) and a left-right difference of $\leqslant 20\,\mathrm{dB}$ at the individual octave frequencies between 0.5 and 4 kHz. The hearing-impaired subjects were stratified into two groups: subjects in the mild loss (ML) group ($n = 19$) had a $PTA_{(0.5,1,2,4\,\mathrm{kHz})} \leqslant 40\,\mathrm{dB\,HL}$ and subjects in the severe loss (SL) group ($n = 21$) had a $PTA_{(0.5,1,2,4\,kHz)} > 40\,dB\,HL$. The average $PTA_{(0.5,1,2,4\,kHz)}$ was 33 dB (±5 dB) in the ML group and 57 dB (±11 dB) in the SL group.

Twenty-one subjects were included in the normal-hearing (NH) reference group. These subjects were tested with a simulated unilateral (conductive) hearing loss in order to obtain information about the maximum possible bilateral benefit. They had a mean age of 27 years (range: 20–40 years) and a $PTA_{(0.5,1,2,4\,kHz)} < 20\,dB$ HL. The age difference between the centers and the groups will be addressed in the Discussion section.

All subjects were recruited via posters or approached at the local clinic for participation. They gave written informed consent and received compensation for participating. Approval for the project (NL32577.018.10) was given by the Ethical Review Board (METC AMC).

HA Fittings

Sixteen of the 19 subjects with a mild hearing loss used an open canal fit. They used a dome which left the ear canal almost entirely open; 19 of the 21 subjects with moderate to severe hearing loss used a custom earmould with a venting diameter between 0 and 3 mm. Overall, all subjects but one wore behind-the-ear HAs. All subjects had the same HAs in both ears.

Insertion gain (IG) measurements were conducted for all hearing-impaired subjects using the International Speech Test Signal (ISTS) at 65 dB SPL (Holube, Fredelake, Vlaming, & Kollmeier, 2010). The ISTS is a non-intelligible speech signal, created by segmenting and mixing running speech in six different languages. It is shaped according to the long-term average speech spectrum (Byrne et al., 1994). The HA settings to which the subjects were accustomed were used in order to represent their daily life situation. This choice led to a strong face validity of the study, but possible heterogeneity within the study group. This aspect will be addressed further in the Discussion section. All HAs used compressive gain. The goodness of HA fit (Byrne et al., 1992) was specified as the root mean square value of the difference between the measured IG values at 0.5, 1, and 2 kHz for a 65-dB input signals and the target values based on the National Acoustic Laboratories (NAL)-RP method (Dillon, 2012).

Measurements

All laboratory measurements were conducted in a sound attenuating, anechoic room. The reverberation time (T_{30}) was below 0.13 seconds for the frequencies between .25 and 4 kHz in both centers. Loudspeakers with a flat frequency response between 0.1 and 18 kHz (±3 dB) were used and were calibrated at the position of the subject's head using stationary speech-shaped noise. The distance between the loudspeakers and the subject was smaller than the critical distance in both centers. As a result, all measurements were done in the direct sound field, such that the influence of the room acoustics was minimal. Testing of the subjects in the ML and SL groups was done with one HA (unilateral condition) or with two HAs (bilateral condition). The aided ear in the unilateral condition was the ear with the better $PTA_{(0.5,\ 1,\ 2,\ 4\,kHz)}$. In 27 subjects, the interaural difference was $\leqslant 3$ dB.

In the NH group, a moderate unilateral conductive hearing loss was simulated by blocking one ear using a foam earplug and an earmuff in order to guarantee sufficient attenuation (Butler, 1986). This was always the ear with the poorer $PTA_{(0.5,1,2,4\,kHz)}$. This group was included in order to investigate the maximum possible effect on the different tests. The limitations of comparing a simulated conductive hearing loss with a true sensorineural hearing loss will be addressed in the Discussion section.

All tests using speech materials were conducted with VU98 sentences (Versfeld, Daalder, Festen, & Houtgast, 2000) at the AMC, and Oldenburger Satztest (Oldenburger Sentence Test; OLSA) sentences (Wagener, Brand, & Kollmeier, 1999) at the HZO. The sequence of tests was balanced and pseudo-randomized using Latin squares (Wagenaar, 1969). All subjects were instructed not to move their head during the experiments. Instructions were repeated if necessary.

Speech reception in noise. Speech reception in noise was assessed using the SRT with fixed or switching sources. In all cases, an adaptive up-down procedure was used to estimate the SNR at which 50% of the sentences were correctly repeated (Plomp & Mimpen, 1979). When testing with fixed sources, speech was presented from the front (0°) and stationary noise, spectrally matched with the speech, at regular conversation level (65 dB) was presented from either 0°, the unilaterally aided side at +90°, or the unilaterally unaided side at −90°. Consequently, a total of three loudspeaker configurations were used (van Esch et al., 2013). The ear which was unaided in the unilateral condition varied between subjects. For ease of reading, the unilaterally aided side is always indicated with a positive angle (+) and the unilaterally unaided side with a negative angle (−).

Speech reception tests were also conducted using switching sources. In this test, two lists of sentences, corresponding to different spatial conditions, were measured during one run. Each next presentation was selected randomly from one of the two spatial conditions: one condition with the speech signal from the loudspeaker at −45° and one with the speech signal from the loudspeaker at +45°. The ISTS was used as a masking signal and was presented from the opposite loudspeaker (at +45° or –45°, respectively). Within each list, an adaptive procedure was applied, as in regular SRT tests. This resulted in two SRT values: one for each spatial condition. The subject had no prior knowledge about the direction of the speech and noise (Boymans et al., 2008; Goverts, 2004). For all loudspeaker configurations, the difference in SRT between the bilateral and unilateral condition is referred to as the Bilateral SRT Benefit.

Listening effort. Listening effort was determined using a categorical scaling procedure by presenting speech at 65 dB SPL from the unilaterally unaided side (−90°) and the ISTS from the unilaterally aided side (+90°). This was done at five different SNRs, roughly at a range around the SRT in stationary noise (group ML: −9, −6, −3, 0, and 3 dB; group SL: −3, 0, 3, 6, and 9 dB; group NH: −12, −9, −6, −3, and 0 dB). Each SNR was used twice per condition, and the order was randomized. The subjects were asked to indicate on a touchscreen how much effort it took to listen to the speech. A 7-point scale with 0.5 intervals was employed, ranging from 0 (*no effort*) to 6 (*extreme effort*; Luts et al., 2010).

Acceptable noise level. The ANL was assessed by determining the level of uncorrelated International Collegium of Rehabilitive Audiology (ICRA-1) noise (Dreschler, Verschuure, Ludvigsen, & Westermann, 2001), presented from +90° and −90° (energetically summed), that was acceptable when listening to speech presented from 0°. The subjects were asked to determine the ANL in six steps by controlling the sound level of the speech and noise via buttons on a touchscreen. The first three steps involved adjusting speech in quiet to a comfortable level, whereas the next three steps involved setting the background noise to an acceptable level with the speech level fixed. The setup and instructions were as described in Nabelek, Tucker, and Letowski (1991) and Freyaldenhoven et al. (2006). The ANL was assessed twice per condition.

Localization. The setup for localization consisted of eight loudspeakers evenly distributed over 360° in the azimuthal plane (i.e., 45° apart). The test was adapted from Goverts (2004) and Boymans et al. (2008) and was chosen because of the realistic test environment using daily life sounds and unexpected timing of the target sound.

The subjects' localization abilities were assessed by asking them to identify which loudspeaker produced a telephone bell sound presented at 65 dB(A) (with roving between −5 and +5 dB in 1 dB steps to reduce intensity cues). Most of the energy of the target signal was concentrated between 1 and 4 kHz. The target sound was presented from one of five loudspeakers in the frontal plane between 4 and 10 seconds after the previous answer. Every 0.7 seconds a new, randomly chosen daily life background sound (a church bell, a crying baby, a chirping bird, water being poured out of a bottle, a guitar, a barking dog, or a siren) was presented from one of the other seven loudspeakers at 65 dB(A). The duration of each sound varied between 2.2 and 3.7 seconds and consequently, after the sound field was built up, three to five different background sounds always played simultaneously. Subjects were asked to point to the loudspeaker of their choice and were instructed not to move their heads. The experimenter repeated this instruction when head movements were observed. In both the unilateral and bilateral condition, the target sound was presented six times from each loudspeaker. The total RMS error was calculated.

Spatial detection. The localization array was also used to assess spatial detection. In the same sound field as described earlier (with daily life background sounds), the ascending method of limits was used as a first order approximation to estimate the detection threshold. The target signal was presented repeatedly from one of the five frontal loudspeakers with an increasing level of 2 dB per presentation. Subjects had to raise their hand when detecting the target signal. The observer logged the level at the moment of detection as the spatial detection threshold for that specific direction. After detecting the signal, another loudspeaker was chosen randomly, and the procedure was repeated. For every condition, the detection thresholds were determined three times for each loudspeaker in the frontal plane.

Statistics

The sample size per group was based on the speech reception in noise data of Boymans et al. (2008). Assuming a bilateral benefit of 0.4 dB (Figure 2 in Boymans et al., 2008) using a female speaker from the unilaterally aided side, a minimum sample size of 20 is needed to detect an effect (with a power of 80% and $\alpha = 0.05$).

Results were analyzed with a one-way repeated measures analysis of variance (ANOVA), using SPSS (SPSS version 20.0.0). The use of one or two HAs was incorporated as the within-subjects factor. Group (severity of hearing loss) and Centre were incorporated as between-subjects factors. A post hoc ANOVA was performed to test the effect of the second HA per group. No further statistical analysis was conducted for the NH group, since this group merely served as a reference group. For the LES, the effect of the second HA was tested using the nonparametric Wilcoxon rank test. Critical values were corrected using a Bonferroni correction based on the total number of outcome measures.

Effect size. To evaluate the effect of the different tests within the test battery, the effect size (r) was calculated. For parametric tests, the F-statistic of the one-way repeated measures ANOVA was used (Field, 2009b):

$$r = \sqrt{\frac{F(1, df_R)}{F(1, df_R) + df_R}}$$

For LES (non-parametric), the Z-score from the Wilcoxon rank test was used (Field, 2009a):

$$r = \frac{Z}{\sqrt{N}}$$

N represents number of subjects and df_R represents degrees of freedom. The absolute value of r lies between 0 (no effect) and 1 (maximum effect).

Results

In all statistical models, the slope of the audiogram between 0.5 and 4 kHz and the goodness of HA fit were introduced as covariates. These factors did not contribute significantly to any of the outcome measures and were, therefore, excluded from the remainder of the analyses. Results of all outcome measures are depicted in Table 1. For all outcome measures, a negative difference corresponds to a bilateral benefit.

Speech Reception in Noise

Figure 1 shows results for the speech reception tests with fixed and switching sources. No significant effects were observed when speech and noise were presented from the front (S_0N_0). When accounting for the test-variability, (all values within twice the standard deviation were considered equal) one subject in the ML group had a bilateral disadvantage and none in the other groups.

When presenting noise from the unilateral unaided side (negative angle), no bilateral benefit was seen with either fixed sources, $F(1,36) = 3.4$, $p > .05$, or switching sources, $F(1,36) = 0.6$, $p > .05$. In other words, adding a second HA at the ear closest to the noise source does not have a positive or negative effect on the SRT. Also, no interaction between a second HA and group was seen for either setup.

There was a significant bilateral benefit when presenting the noise from the unilaterally aided side (positive angle) in both setups (fixed sources: $F(1,36) = 27.1$, $p < .001$; switching sources: $F(1,36) = 35.9$, $p < .001$. The magnitude of this effect increased with increasing hearing loss (fixed sources: $F(1,36) = 16.5$, $p < .01$; switching sources: $F(1,36) = 35.9$, $p < .001$, for the interaction between the second HA and group). For either setup, post hoc analysis showed no significant effect for the second HA for subjects in group ML (fixed sources: $F(1,17) = 1.4$, $p > .05$; switching sources: $F(1,17) = 1.9$, $p > .05$). However, in group SL, there was a bilateral benefit when using fixed sources (−4.1 dB: $F(1,19) = 30.4$, $p < .001$) and switching sources (−6.4 dB: $F(1,19) = 40.5$, $p < .001$). Based on the results of group NH, the maximum benefit is −7.8 dB using fixed sources and −12.8 dB using switching sources.

The above results suggest that adding a second HA increases performance, but only for hearing losses larger than 40 dB ($PTA_{(0.5,1,2,4\,kHz)}$) when the unaided ear is not able to compensate for the head shadow effect.

Significant center effects were found for all speech reception tests. The 50%-point at the HZO was generally lower than at the AMC, which is in correspondence with the normative data for the German and Dutch speech material (Versfeld et al., 2000; Wagener et al., 1999). No interaction effects with regard to the center were found.

Listening Effort

Figure 2 presents the results for listening effort at −3 dB and 0 dB. These were the conditions common to all subjects. Statistical analyses were performed using a Wilcoxon rank test and a bilateral benefit was seen for all groups at an SNR of −3 dB (group ML: $Z = -3.3$, $p < .05$; group SL: $Z = -3.4$, $p < .05$). The magnitude of the effect was 0.5 LES units in group ML and 2 LES units in group SL. Only the severely hearing-impaired subjects benefited from the second HA at an SNR of 0 dB (group ML: $Z = -0.7$, $p > .05$; group SL: $Z = -3.8$, $p < .01$). When analyzing the different test conditions combined a median bilateral benefit of 0.5 points is present in group ML ($Z = -4.9$, $p < .001$) and of 1.3 units in group SL ($Z = -7.4$, $p < .001$). In group NH, this difference was 2.3 units (not tested for significance).

A median benefit of 0.5 points on a 7-point scale is relatively small. A total of 74% of the subjects in group ML indicated that speech reception took less effort with two HAs than with one HA, but none of the difference scores were above two points.

ANL

No effect of a second HA was found for ANL, $F(1, 36) = 0.5$, $p > .05$, nor were there any interactions with group or center. Therefore, no post hoc tests were done. A center effect was present, $F(1, 36) = 20.6$, $p < .01$. Results are depicted in Figure 3.

Localization

In the results presented in Figure 4, the RMS error was averaged over all angles. A significant effect was found for

Table 1. All Outcomes of the Laboratory Tests and Questionnaires.

			1HA						2HA						Difference (2HA–1HA)					
		Effect secondnd	ML group		SL group		NH group		ML group		SL group		NH group		ML group		SL group		NH group	
Test	Condition	HA	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
SRT Fxd [dB]	S0°N0°	*nt*	−3.1	2.1	−0.4	3.6	−5.2	2.6	−2.8	2.6	−1.2	3.4	−6.1	2.5	0.3	1.2	−0.7	1.4	−0.9	1.2
	S0°N−90°	ns	−8.0	2.9	−3.5	4.4	−9.7	4.6	−7.9	3.0	−4.3	4.0	−11.9	3.1	0.1	0.9	−0.9*	1.6	−2.2	2.4
	S0°N+90°	***	−7.0	3.2	−0.6	4.4	−3.6	2.5	−7.5	3.3	−4.7	4.0	−11.3	3.3	−0.5	1.9	−4.1***	3.4	−7.8	2.0
SRT Sw [dB]	S45°N−45°	ns	−15.3	3.4	−5.3	5.8	−21.6	4.9	−15.4	3.2	−6.1	6.4	−24.3	4.8	−0.1	2.3	−0.8	4.4	−2.7	2.3
	S−45°N+45°	***	−14.6	5.5	−2.0	6.0	−11.3	3.7	−15.4	3.6	−8.4	4.8	−24.1	4.5	−0.7	3.6	−6.4	4.6	−12.8	2.6
LES [LES units]	−12 dB SNR	*nt*					5.8	−0.3 0.3					4.0	−1.1 0.3					−1.8	−0.6 0.3
	−9 dB SNR	*nt*	4.5	−0.5 0.6			5.3	−0.8 0.5	4.1	−0.8 0.9			3.3	−1.3 0.3	−0.5	0.0 0.5			−2.5	−0.3 1.0
	−6 dB SNR	*nt*	3.5	−0.5 0.6			4.8	−0.6 0.8	3.4	−0.5 0.5			2.5	−1.3 0.8	0.5	−0.1 0.4			−2.3	−1.1 0.5
	−3 dB SNR	***	3.0	−1.0 1.0	5.0	−1.0 1.0	4.0	−0.8 0.8	2.5	−1.0 0.5	3.0	−0.5 1.5	1.3	−0.8 1.3	−0.5*	−0.3 0.3	−2.0*	−0.6 1.5	−2.8	−0.3 1.5
	0 dB SNR	**	2.0	−0.5 0.5	4.5	−1.0 0.5	3.0	−1.1 1.3	1.8	−0.8 1.0	2.5	−0.8 0.8	1.5	−1.5 0.6	−0.4	−0.1 0.6	−1.5**	−1.3 1.0	−1.8	−0.8 0.8
	3 dB SNR	*nt*	1.5	−1.0 0.5	3.5	−0.8 1.1			0.8	−0.8 0.8	2.0	−1.8 0.6			−0.5	−0.5 0.5	−1.5	−1.3 1.1		
	6 dB SNR	*nt*			2.5	−1.0 0.6					0.6	−0.6 0.6					−1.3	−0.6 0.8		
	9 dB SNR	*nt*			1.4	−0.4 0.8					0.0	0.0 1.0					−1.0	−0.5 1.0		
ANL [dB]		ns	8.7	8.5	5.1	5.3	7.0	5.4	7.3	8.7	5.4	6.5	5.7	6.0	−1.3	2.7	0.3	5.4	−1.4	4.0
Det [deg]	Target −90°	**	47.2	4.8	56.6	5.7	44.2	5.0	46.7	4.3	51.4	5.6	38.5	1.9	−0.3	3.3	−5.2***	4.8	−5.7	4.3
	Target −45°	*nt*	48.3	2.8	55.4	4.7	44.0	4.2	47.7	2.7	52.6	5.3	38.1	2.1	−0.6	2.9	−2.8	4.8	−5.9	4.2
	Target 0°	*nt*	48.3	2.4	52.6	4.4	40.1	3.5	47.7	3.7	51.4	6.4	38.5	2.2	−0.6	4.0	−1.2	4.4	−1.6	3.6
	Target 45°	*nt*	49.4	2.8	51.4	4.4	38.8	2.2	46.8	3.7	52.1	7.1	38.7	2.5	−2.6	4.0	0.7	5.2	−0.1	2.7
	Target 90°	*nt*	47.8	2.9	51.0	4.7	38.7	2.0	46.4	3.5	51.0	6.2	38.6	1.8	−1.6	3.0	0.1	3.9	−0.1	2.2
	All angles	ns	48.2	2.4	53.4	4.2	41.2	2.5	47.1	2.9	51.7	5.5	38.5	1.4	−1.2	2.2	−1.7	3.4	−2.7	2.0
Loc [deg]	Target −90°	*nt*	28.6	25.3	75.7	46.3	75.9	36.1	21.6	18.2	28.5	18.6	2.6	12.0	−6.9	25.2	−47.2	36.8	−73.3	37.7
	Target −45°	*nt*	20.4	16.0	49.3	28.1	64.3	24.1	18.8	15.0	33.1	13.4	1.7	5.4	−1.6	13.5	−16.2	25.9	−62.6	24.3
	Target 0°	*nt*	17.6	18.0	47.5	28.7	51.1	11.5	6.0	12.9	25.2	19.7	0.0	0.0	−11.6	20.4	−22.3	22.1	−51.1	11.5
	Target 45°	*nt*	20.5	13.2	36.1	13.5	28.9	13.2	20.3	23.6	33.6	10.8	1.7	5.4	−0.2	22.5	−2.5	17.0	−27.1	13.9
	Target 90°	*nt*	15.2	18.9	36.0	22.0	54.7	19.4	17.6	25.4	35.5	20.6	5.0	11.2	2.4	20.7	−0.6	26.4	−49.7	22.7
	All angles	***	25.6	11.0	55.5	19.4	60.0	13.0	22.3	14.2	34.5	8.6	3.2	8.0	−3.4	14.4	−21.0***	18.1	−56.8	15.6

Note. Means (m) and *SD* are given for the different groups with one and with two hearing aids. For LES, the median and the upper and lower quartile are shown. In the last column, the difference between one and two hearing aids is depicted. The overall effect of two hearings aids is either not tested (nt), not significant (ns) or significant with $^{*}p < .05$, $^{**}p < .01$, or with $^{***}p < .001$. Per group the difference between one and two hearing aids is depicted in the final column. Bonferroni's method was used to correct for the number of comparisons.

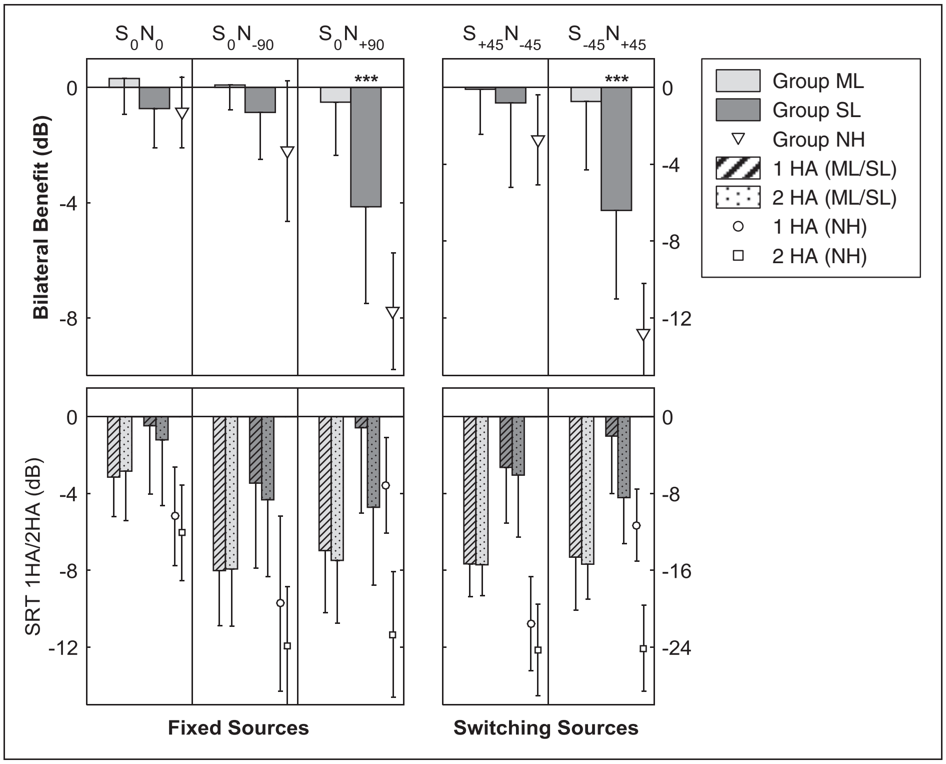

Figure 1. Speech reception in noise. The top row shows the mean bilateral benefit in SRT: $*p < .05$, $**p < .01$, or $***p < .001$. A negative value represents an advantage. The bottom row shows the mean SRTs with one or with two hearing aids. The panels on the left show the results with continuous noise presented from fixed sources. The panels on the right show the results with the ISTS presented from switching sources. The whiskers represent the standard deviation. Negative angles correspond to the unilaterally unaided side.

the second HA, $F(1, 36) = 22.0$, $p < .001$, as was an interaction between HA and group, $F(1, 36) = 11.2$, $p < .05$. Post hoc testing showed a clear bilateral advantage for group SL of $-21°$, $F(1, 19) = 27.9$, $p < .001$, but not for subjects in the ML group, $F(1, 17) = 1.1$, $p > .05$. Based on the results of group NH, the maximum benefit is $-57°$.

Figure 5 shows the same data, plotted as individual data points. The left panel shows the bilateral benefit as a function of low-frequency hearing loss (this average was chosen in order to compare the results to those of Byrne et al., 1992. See the Discussion section for this comparison). The right panel shows the RMS error in the unilateral and bilateral conditions. In both plots, the least squares fit of the data is plotted. In the right panel, it can be seen that the RMS error increases with hearing loss in both the unilateral and bilateral condition, but that the slope is larger with one HA (0.92°/dB vs. 0.33°/dB). As a consequence, the bilateral benefit increases with increasing hearing loss at a rate of 0.58°/dB as can be seen in the left panel.

Spatial Detection

Although a trend toward greater bilateral benefit in subjects with larger hearing loss was observed for spatial detection, no significant effect of second HA use was found when looking at the average spatial detection threshold, $F(1, 36) = 9.8$, $p > .05$. See also Figure 6.

Analyzing only the spatial detection threshold when the target sound was presented from the unilaterally unaided side (at $-90°$), an interaction between the second HA and group was seen, $F(1, 36) = 20.4$, $p < .01$. Post hoc testing showed that, for this angle, the second HA has a significant effect in group SL, $F(1, 19) = 27.2$, $p < .001$, but not in group ML, $F(1, 17) = 0.2$, $p > .05$. In Figure 7, the data per angle are presented in a polar plot. Here, it can be seen that presentation of the target signal from the unilateral unaided side ($-90°$) leads to differences between performance with one and with two HAs in groups SL and NH.

Effect size

According to Cohen (1992) effect sizes of .1, .3, and .5 represent a small, medium, and large effect, respectively. In Figure 8, the effect sizes for all tests are depicted and sorted based on the magnitude of the effect in group NH, which represents the maximum bilateral benefit. In this group, the SRT with switching sources ($S_{+45°}N_{-45°}$) give the largest bilateral benefit and the ANL the smallest. In

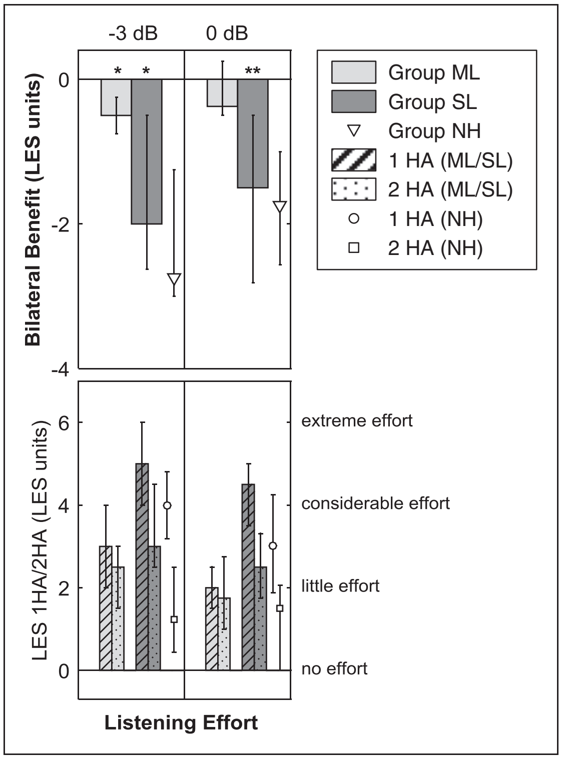

Figure 2. Listening effort scaling. The top panel shows the median bilateral benefit of the LES: $*p<.05$, $**p<.01$, or $***p<.001$. A negative value represents an advantage. The bottom panel shows the median LES with one or with two hearing aids. The whiskers represent the interquartile range.

both group NH and group SL, the largest benefit is seen using the speech reception tests with noise from the unilaterally aided side, localization, and listening effort. However, the effect sizes for speech reception and localization drop drastically in group ML, although the effect size of listening effort remains large.

Discussion

The aim of this study was to investigate the effect of bilateral HAs in subjects with mild and moderate-to-severe hearing loss in different dimensions of performance. For the severely hearing-impaired subjects (group SL), addition of the second HA was found to have a significant effect on localization, listening effort and on speech reception in noise when noise was presented from the unilaterally aided side (positive angles). For mildly hearing-impaired subjects (group ML), only listening effort revealed a significant bilateral benefit.

Speech Reception in Noise

One can only experience a bilateral benefit when there is room for improvement in the unilateral condition. Figure 1 shows that speech reception using switching

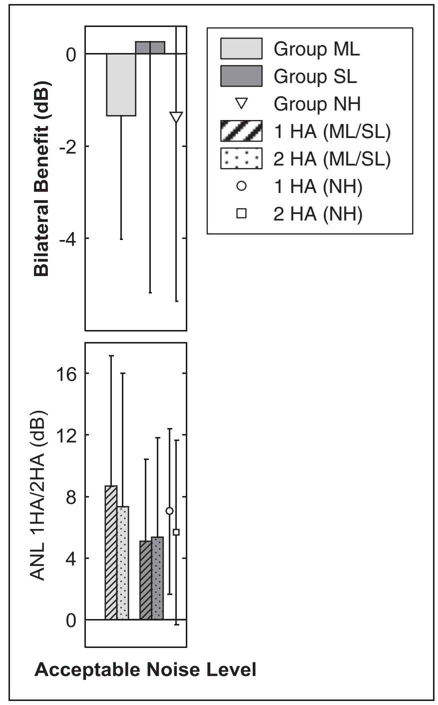

Figure 3. Acceptable noise level. The top panel shows the mean bilateral benefit of the ANL: $*p<.05$, $**p<.01$, or $***p<.001$. A negative value represents an advantage. The bottom panel shows the mean ANL with one or with two hearing aids. The whiskers represent the standard deviation.

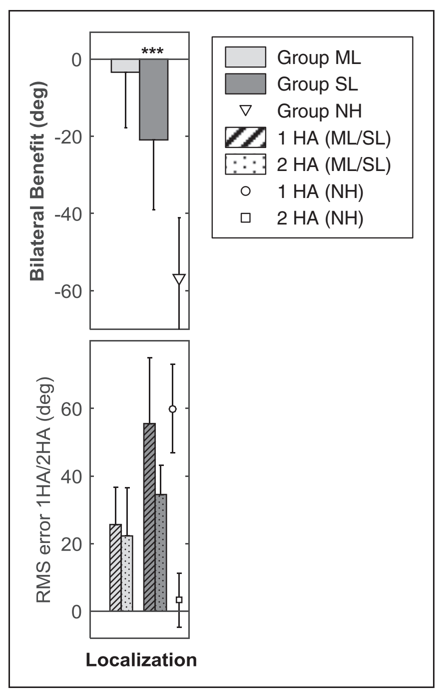

Figure 4. Localization. The top panel shows the mean bilateral benefit in RMS error: $*p<.05$, $**p<.01$, or $***p<.001$. A negative value represents an advantage. The bottom panel shows the mean RMS error with one or with two hearing aids. The whiskers represent the standard deviation.

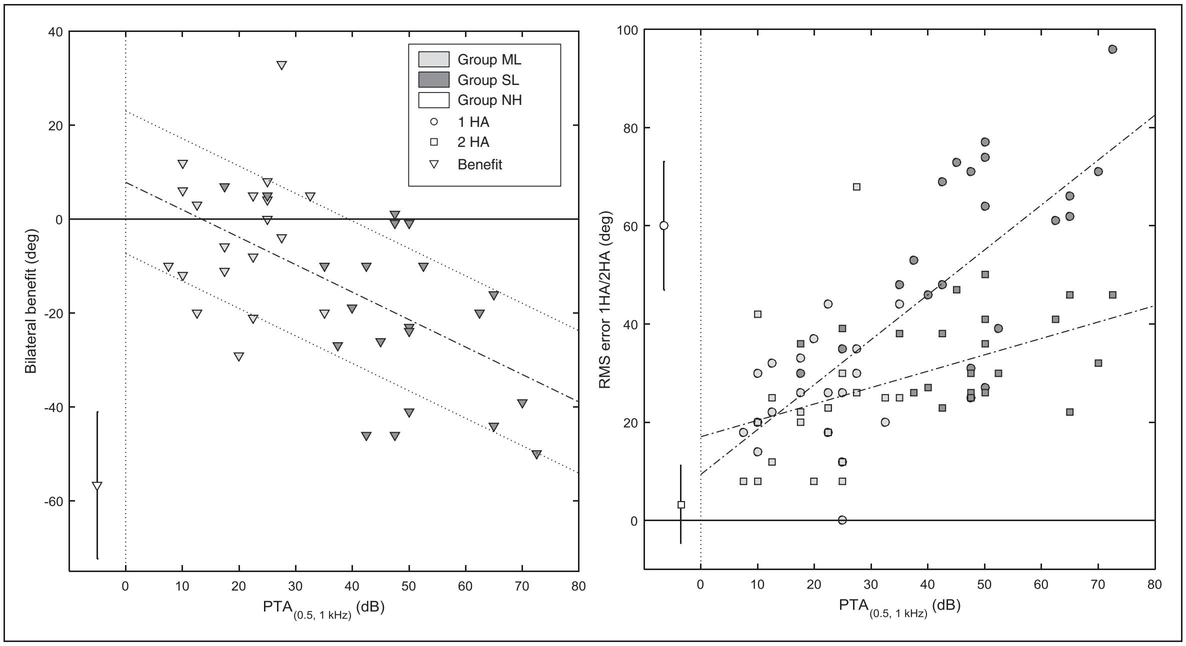

Figure 5. Localization. The figure shows the individual data with 1 and with 2 HAs (left panel) and the bilateral benefit (right panel) averaged over all angles, plotted against low frequency hearing loss. The dash-dot lines represent the linear least squares fit of the unilateral data ($y = 0.92x + 9.36$; $r^2 = 0.58$; $p < .001$), the bilateral data ($y = 0.33x + 17.04$; $r^2 = 0.22$; $p < .01$) and the bilateral benefit ($y = -0.58x + 7.80$; $r^2 = 0.33$; $p < .01$). The dashed lines in the right panel represent the standard deviation.

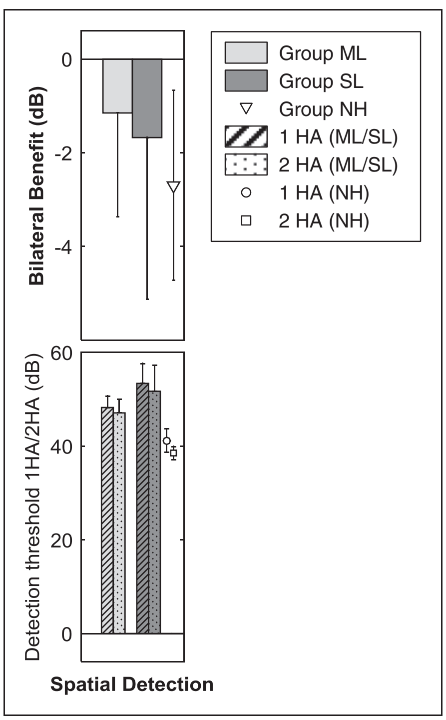

Figure 6. Spatial detection. The top panel shows the mean bilateral benefit of the detection threshold, averaged over all angles: *$p < .05$, **$p < .01$, or ***$p < .001$. A negative value represents an advantage. The bottom panel shows the mean detection threshold with one or with two hearing aids. The whiskers represent the standard deviation.

sources with one HA deteriorates by only 0.7 dB ($p > .05$) when moving the noise source from the unaided to the aided side (the difference between $S_{+45}N_{-45}$ and $S_{-45}N_{+45}$). In the SL group, this deterioration is 3.3 dB ($p < .01$) and in the NH group (with simulated unilateral conductive hearing loss), the deterioration is 10.3 dB ($p < .001$). Similar numbers were observed using the setup with fixed sources. Obviously, the location of the noise source has no strong effect on speech reception with one HA in subjects with a mild hearing impairment, which is illustrated in the bottom panels. In other words: due to the low hearing thresholds of these subjects ($PTA_{.5/1/2/4\,kHz} < 40$ dB HL), unaided performance is relatively good, which means there is only little room for improvement. Similar results were also found by Festen and Plomp (1986).

When noise is presented on the unilaterally unaided side (S_0N_{-90} and $S_{+45}N_{-45}$), it is expected that binaural unmasking plays a role when adding a second HA. Figure 1 shows that hearing-impaired subjects received little or no benefit from their second HA using this setup. Binaural unmasking resulted in an average improvement of 2 dB in the normally hearing subjects (triangle), which is less than what Marrone, Mason, and Kidd (2008) found. They reported a binaural benefit of 8 to 12 dB for decreasing reverberation times using a speech masker. Our results correspond to the results of Markides (1979), who reported a benefit of 2–3 dB when

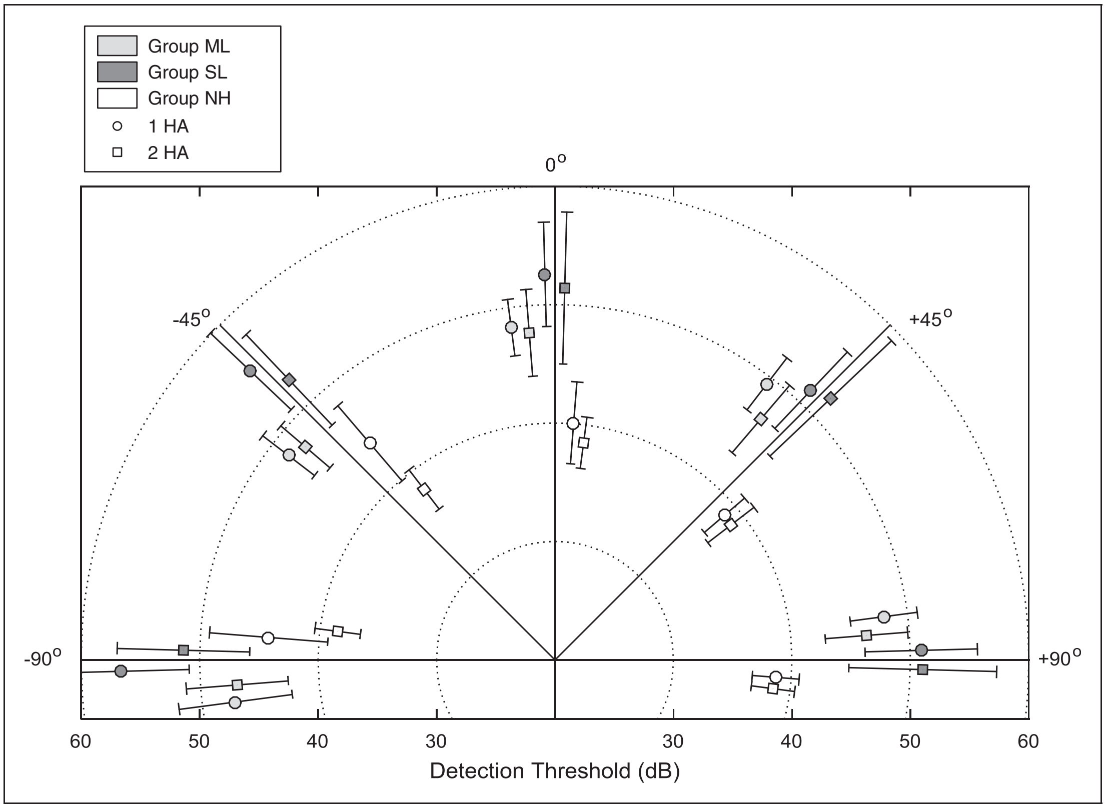

Figure 7. Spatial detection. Polar plot of the detection threshold per angle in the frontal plane. Negative angles correspond to the unilaterally unaided side.

presenting word lists in stationary noise. However, Markides found his results when presenting speech from $+45^{\circ}$ and noise from -45°, which is different from our fixed setup. In the current study, the effect of binaural unmasking for both the ML and SL groups was less than 1 dB. Festen and Plomp (1986) stated that five subjects in their study with a PTA (0.5, 1, 2 kHz) > 60 dB HL clearly profited from a second HA contralateral to the noise source. Three of these subjects also appeared to benefit from an additional HA ipsilateral to the noise source. Their setup was similar to that of the current study, but they used higher noise levels. In the current study, 6 of 21 subjects in group SL had a PTA (0.5, 1, 2, 4 kHz) > 60 dB HL, but only the two subjects with the highest thresholds (>70 dB HL) appeared to benefit from binaural unmasking.

Noble and Gatehouse (2006) discussed the importance of bilateral fittings in demanding listening situations and with rapidly dividing attention. In the current study, we attempted to create a more dynamic environment by presenting noise and speech using switching sources. However, no convincing differences in effect size between switching and fixed source tests were observed (see Figure 8). A possible explanation for this finding is that the experiment was still relatively static in comparison with daily life situations. A combined speech reception in noise and localization task with more than two sources might be a better approximation of the dynamic situations described by Noble and Gatehouse.

Henkin et al. (2007) tested 28 mild to moderately hearing-impaired subjects with a mean age of 73 years with monosyllabic words at +10 dB SNR using diotic stimulation. The authors found 71% of the subjects to have a bilateral disadvantage, which they attributed to binaural interference. It must be noted that the authors did not take the test variability into account and considered all difference scores larger than 0% to be a bilateral disadvantage (with 0.18% being the smallest difference). Walden and Walden (2005) reported similar findings when using one loudspeaker from the front using sentences in four talker babble noise. Over 82% of subjects performed better with unilateral than with bilateral amplification. It is unclear what criterion was used by the authors to classify a bilateral disadvantage. McArdle et al. (2012) reproduced the study by Walden and Walden with a similar study group and found that only 20% of subjects had poorer performance with a bilateral fit. In the current study, 35% of all hearing-impaired subjects experienced bilateral disadvantage with the criterion set to the unrealistic value of 0 dB, similar to the criterion used by Henkin et al. (2007). When setting the criterion to twice the standard

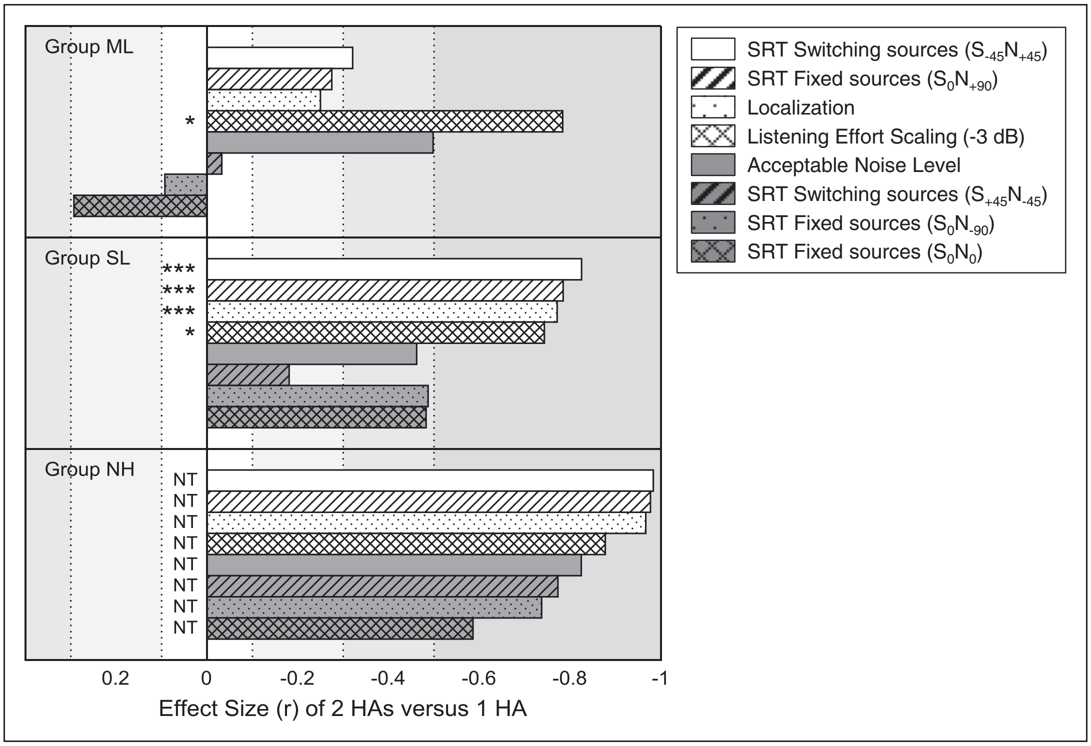

Figure 8. Effect size for all tests in the test battery, calculated according to Field (2009a, 2009b): NT (not tested), $*p < .05$), $**p < .01$, or $***p < .001$. An effect size of .10, .30, and .50 corresponds to a small, medium, and large effect, respectively (Cohen, 1992). A negative score (in this plot on the right side) implies a bilateral benefit. Data are sorted based on the results of the normal hearing reference group (group NH). Negative angles correspond to the unilaterally unaided side.

deviation, 2.5% experienced a bilateral disadvantage (one subject in group ML) and 8% experienced a bilateral advantage (three subjects in group SL). No effect of age was seen. Our results suggest that hearing-impaired subjects receive only little benefit from their second HA with colocated sources, but also no disadvantage. This disagrees with the findings of Henkin et al. and Walden and Walden but is in agreement with the results of McArdle et al.

Listening Effort Scaling

Noble (2006) stated that the benefit of two HAs lies in reduced listening effort. This statement is based on self-report measures and was confirmed in our laboratory. LES at an SNR of −3 dB showed that both hearing-impaired groups experienced significant benefit from their second HA. LES is the only test that demonstrated a bilateral benefit in mildly hearing-impaired subjects. The median benefit at −3 dB was −0.5 points (interquartile range −0.75 to −0.25) on a 7-point scale, and although statistically significant, it is unclear whether this difference is clinically relevant.

Sarampalis, Kalluri, Edwards, and Hafter (2009) tried to find an objective measure for listening effort in order to evaluate noise reduction in HAs. At SNRs where the speech was highly intelligible, they found a decreased cognitive effort for increasing SNRs. This decreased cognitive effort at maximum speech intelligibility is in agreement with the results from the current study. It appears that when a ceiling effect is reached for speech intelligibility, there is still room for reducing listening effort.

Rennies, Schepker, Holube, and Kollmeier (2014) assessed listening effort in normally hearing subjects via headphones using the same procedure as was used in the current study. Between SNRs of −10 and +6 dB, they found a consistent change in effort of approximately −0.3 points per dB SNR. For SNRs of −2, +2, and +6 dB their intelligibility scores were 100%. These results are in line with our findings, where the slope ranged between −0.23 and −0.31 points per dB. This implies that the median bilateral benefit we found for subjects with mild hearing loss (0.5 points) corresponds to an increase in SNR of about 2 dB. Such an increase in SNR can be considered as clinically relevant. However, to transform this test into a useful clinical tool remains challenging, since such a small difference is difficult to assess on an individual level. Listening effort itself is an informative outcome, but there is need for further research if we wish to assess bilateral benefit in terms of listening effort, particularly for mildly hearing-impaired subjects.

Acceptable Noise Level

Freyaldenhoven et al. (2006) stated that bilateral amplification probably does not affect acceptance of noise. Their results for unilateral and bilateral conditions are comparable to our results, but the variability in our study is larger (the number of subjects is similar). Brannstrom, Holm, Kastberg, and Olsen (2014) tested a non-semantic version of the ANL in 32 normally hearing subjects aged between 18 and 40 years. They used the ISTS as speech and presented both speech and noise diotically through earphones. The ANL measurement was repeated 12 times and analyzed in four blocks of three tests. Based on their data, the authors suggest that, in order to obtain an accurate ANL, more training sessions and more replications are needed than originally proposed by Nabelek et al. (1991). This implies that the number of sessions in the current study may have been too small to obtain an accurate ANL. In their discussion article about the ANL, Olsen and Brannstrom (2014) reviewed the coefficient of repeatability (CR): in repeated measurements, 95% of the absolute differences between two repeated measures will be lower than the CR. The CR values in the studies examined by the authors ranged from 4.7 to 14 dB. In the current study, all but one of the subjects showed a bilateral benefit between −10 and +10 dB. Given the fact that this is below some of the CR values reported by Olsen and Brännström it is possible that, even if a bilateral benefit was present, it cannot be demonstrated using the current setup.

Localization

No bilateral benefit for localization was seen for the mildly hearing-impaired subjects as opposed to the subjects in group SL. Byrne et al. (1992) stated that localization in the horizontal plane is mainly affected by low frequency sensorineural hearing loss (below 1,500 Hz), but only when this loss reaches 50 dB HL. In the current study, no subjects in group ML had a low frequency threshold >50 dB HL ($PTA_{(0.5,1\,kHz)}$), whereas 29% of the subjects in group SL had a hearing loss of at least 50 dB HL at this frequency (taking 40 dB HL as a threshold, the percentages are 0% and 76% for groups ML and SL, respectively). Most subjects with relatively good low frequency hearing had an open HA fit. In this type of fitting, low frequency sounds enter the ear via the direct acoustic path. Consequently, the low frequency difference at the eardrum between the unaided and aided condition is smaller for open fittings than for fittings with a custom earmould. This might partially explain that subjects in group ML performed relatively well in the unilateral condition and did not benefit from the extra amplification in the bilateral condition. Furthermore, compressive HAs with short attack times can cause deterioration of localization cues in the higher frequency regions (Musa-Shufani, Walger, von Wedel, & Meister, 2006), which also may have contributed to the lack of bilateral benefit for this group.

One aspect of the laboratory setup that could be improved is the resolution of the test. The target signal was presented from one of five loudspeakers, separated by an angle of 45°. The task gets more difficult when the angle between loudspeakers gets smaller, creating a more challenging task even for mildly impaired subjects.

Spatial Detection

The localization setup was also used to test spatial detection. The expected bilateral benefit in this experiment is mainly the result of the head shadow effect, occurring predominantly when the target signal is played from the unilaterally unaided side. According to Shaw (1974), using a sound source at +90°, the head shadow effect results in attenuation at the far ear of 6–16 dB between 700 and 5,600 Hz. The sound energy of our target signal was concentrated in this frequency region. Although the variability was large, on average, the NH subjects displayed a bilateral benefit of around 6 dB with the target signal is presented from the unilaterally unaided side (see Figure 7). The subjects in group SL displayed greater variability, but still a significant bilateral benefit of 4 dB was observed for that angle.

Limitations of this Study

It is challenging to conduct a multicenter study using different test locations and different test leaders. An important difference between centers is the speech material. Not only is the language different but also the structure of the sentences: OLSA sentences are constructed to have the same syntax, whereas the Dutch sentences vary in syntax. Slopes of the performance-intensity functions differed between speech materials. The significant center effects that were found in the analysis are mainly due to these differences. However, no interaction effects with regard to center were observed, implying that the bilateral benefit is similar between the centers. We can, therefore, compare these results without problems.

The intention of this study was to evaluate the performance with HA settings that the subjects were accustomed to. However, IG measurements showed that various HAs provided less amplification than prescribed by the NAL-RP rule (averaged over four frequencies, the difference was larger than 5 dB in 10% of the subjects). The approach used strengthens the face validity of the results but is a weakness in the study design because it contributes to the heterogeneity of the groups. Altogether, the difference between target and IG was found to have no significant effect on the results.

A normal hearing reference group was used to get an indication of the maximum possible bilateral benefit. The subjects in this group were significantly younger than the hearing-impaired subjects. Besides this, the use of an earmuff and an earplug introduced a unilateral conductive hearing loss of at least 40 dB (Abel & Armstrong, 1992). A conductive hearing loss causes linear attenuation, where a sensorineural hearing loss introduces non-linear effects. This leads to differences in perception of suprathreshold sounds due to differences in loudness growth, spectral resolution, temporal resolution, and other aspects. Direct comparison of these groups is, therefore, not possible. Therefore, group NH was not included in the statistical analyses but merely served to obtain a benchmark. Consequently, the differences between the hearing-impaired groups and group NH did not affect the conclusions of the manuscript.

Finally, the decision to investigate unilateral HA use in subjects who wear bilateral HAs by choice introduced a bias toward bilateral benefit. Our subjects were accustomed to bilateral amplification and were tested in both bilateral and unilateral conditions. This may have led to an overestimation in the bilateral benefit they experienced. However, a within-subject design reduces the error due heterogeneity between the subjects, making the detection of a small effect more likely.

Conclusions

Overall, a bilateral benefit was predominantly observed with respect to speech reception in noise, listening effort, and localization. This effect tended to be larger for the severely than for the mildly hearing-impaired subjects. The subjects with mild hearing loss only showed a significant benefit on LES. Besides this, no significant disadvantage of using a second HA was found in any of the laboratory tests.

For the future, it is important to determine which tests are applicable for clinical use. The assessment of listening effort seems to be most valuable for people with mild hearing impairment. However, before this promising outcome measure can be used on an individual level, more research is needed.

An important next step is to focus on those tests with greatest potential for demonstrating bilateral benefit in daily life. This could be investigated by comparing experienced bilateral HA users with HA users that have a symmetrical hearing loss, but prefer unilateral amplification.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors would like to thank the hearing aid companies that funded this research study: GN ReSound, Oticon, Phonak, Siemens, and Widex.

References

Abel, S. M., & Armstrong, N. M. (1992). The combined sound-attenuation of earplugs and earmuffs. *Applied Acoustics*, *36*(1), 19–30.

Akeroyd, M. A. (2014). An overview of the major phenomena of the localization of sound sources by normal-hearing, hearing-impaired, and aided listeners. *Trends Hear*, *18*, 1–7.

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 Hear*, *29*(6), 930–941.

Brannstrom, K. J., Holm, L., Kastberg, T., & Olsen, S. O. (2014). The acceptable noise level: The effect of repeated measurements. *International Journal of Audiology*, *53*(1), 21–29.

Bronkhorst, A. W., & Plomp, R. (1989). Binaural speech intelligibility in noise for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *86*(4), 1374–1383.

Butler, R. A. (1986). The bandwidth effect on monaural and binaural localization. *Hearing Research*, *21*(1), 67–73.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., . . . Ludvigsen, C. (1994). An international comparison of long-term average speech spectra. *Journal of the Acoustical Society of America*, *96*(4), 2108–2120.

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.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

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. (2012). Prescribing hearing aid amplification. In H. Dillon (Ed.), *Hearing aids* (2nd ed., pp. 286–335). New York, NY: Thieme.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International collegium for rehabilitative audiology. *Audiology*, *40*(3), 148–157.

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. F. (1992). Monaural versus binaural hearing: Ease of listening, word recognition, and attentional effort. *Ear and Hearing*, *13*(2), 80–86.

Field, A. P. (2009a). Non-parametric tests. In A. P. Field (Ed.), *Discovering statistics using SPSS* (3rd ed., pp. 539–583). London, Engalnd: SAGE Publications.

Field, A. P. (2009b). Repeated-measures designs (GLM 4). In A. P. Field (Ed.), *Discovering statistics using SPSS* (3rd ed., pp. 457–538). London, England: SAGE Publications.

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.

Goverts, S. T. (2004). *Assessment of spatial and binaural hearing in hearing impaired listeners*. (Dissertation). VU Amsterdam, Amsterdam.

Grothe, B., Pecka, M., & McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiological Reviews*, *90*(3), 983–1012.

Henkin, Y., Waldman, A., & Kishon-Rabin, L. (2007). The benefits of bilateral versus unilateral amplification for the elderly: Are two always better than one? *Journal of Basic Clinical Physiology and Pharmacology*, *18*(3), 201–216.

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology*, *49*(12), 891–903.

Irving, S., & Moore, D. R. (2011). Training sound localization in normal hearing listeners with and without a unilateral ear plug. *Hearing Research*, *280*(1–2), 100–108.

Kim, J. H., Lee, J. H., & Lee, H. K. (2014). Advantages of binaural amplification to acceptable noise level of directional hearing aid users. *Clinical and Experimental Otorhinolaryngology*, *7*(2), 94–101.

Kobler, S., & Rosenhall, U. (2002). Horizontal localization and speech intelligibility with bilateral and unilateral hearing aid amplification. *International Journal of Audiology*, *41*(7), 395–400.

Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M., Buechler, M., ... Spriet, A. (2010). Multicenter evaluation of signal enhancement algorithms for hearing aids. *The Journal of the Acoustical Society of America*, *127*(3), 1491–1505.

Markides, A. (1979). Issuing 2 hearing-aids for simultaneous use. *Journal of Laryngology and Otology*, *93*(6), 575–587.

Markides, A. (1982). The effectiveness of binaural hearing aids. *Scandinavian Audiology. Supplementum*, *15*, 181–196.

Marrone, N., Mason, C. R., & Kidd, G. (2008). Tuning in the spatial dimension: Evidence from a masked speech identification task. *The Journal of the Acoustical Society of America*, *124*(2), 1146–1158.

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.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, *42*, 135–159.

Moore, B. C., Johnson, J. S., Clark, T. M., & Pluvinage, V. (1992). Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss. *Ear and Hearing*, *13*(5), 349–370.

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.

Musa-Shufani, S., Walger, M., von Wedel, H., & Meister, H. (2006). Influence of dynamic compression on directional hearing in the horizontal plane. *Ear and Hearing*, *27*(3), 279–285.

Nabelek, A. K., Tucker, F. M., & Letowski, T. R. (1991). Toleration of background noises: Relationship with patterns of hearing aid use by elderly persons. *Journal of Speech and Hearing Research*, *34*(3), 679–685.

Noble, W. (2006). Bilateral hearing aids: A review of self-reports of benefit in comparison with unilateral fitting. *International Journal of Audiology*, *45*(Suppl 1): S63–S71.

Noble, W., & Byrne, D. (1991). Auditory localization under conditions of unilateral fitting of different hearing aid systems. *British Journal of Audiology*, *25*(4), 237–250.

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., Ter-Horst, K., & Byrne, D. (1995). Disabilities and handicaps associated with impaired auditory localization. *Journal of the American Academy of Audiology*, *6*(2), 129–140.

Olsen, S. O., & Brannstrom, K. J. (2014). Does the acceptable noise level (ANL) predict hearing-aid use? *International Journal of Audiology*, *53*(1), 2–10.

Persson, P., Harder, H., Arlinger, S., & Magnuson, B. (2001). Speech recognition in background noise: Monaural versus binaural listening conditions in normal-hearing patients. *Otology and Neurotology*, *22*(5), 625–630.

Plomp, R. (1976). Binaural and monaural speech-intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise). *Acustica*, *34*(4), 201–211.

Plomp, R., & Mimpen, A. M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, *18*(1), 43–52.

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.

Rennies, J., Schepker, H., Holube, I., & Kollmeier, B. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *Journal of the Acoustical Society of America*, *136*(5), 2642–2653.

Sarampalis, A., Kalluri, S., Edwards, B., & Hafter, E. (2009). Objective measures of listening effort: Effects of background noise and noise reduction. *Journal of Speech Language and Hearing Research*, *52*(5), 1230–1240.

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.

Shaw, E. A. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *The Journal of the Acoustical Society of America*, *56*(6), 1848–1861.

van Esch, T. E., Kollmeier, B., Vormann, M., Lyzenga, J., Houtgast, T., Hallgren, M., . . . Dreschler, W. A. (2013). Evaluation of the preliminary auditory profile test battery in an international multi-centre study. *International Journal of Audiology*, *52*(5), 305–321.

Vaughan-Jones, R. H., Padgham, N. D., Christmas, H. E., Irwin, J., & Doig, M. A. (1993). One aid or two? – More visits please! *The Journal of Laryngology and Otology*, *107*(4), 329–332.

Versfeld, N. J., Daalder, L., Festen, J. M., & Houtgast, T. (2000). Method for the selection of sentence materials for efficient measurement of the speech reception threshold. *The Journal of the Acoustical Society of America*, *107*(3), 1671–1684.

Wagenaar, W. A. (1969). Note on the construction of Digram-Balanced Latin Squares. *Psychological Bulletin*, *72*(6), 3.

Wagener, K., Brand, T., & Kollmeier, B. (1999). Entwicklung und evaluation eines satztests für die deutsche sprache. III evaluation des oldenburger satztests (Development and evaluation of a sentence test for the German language. III: Evaluation of the Oldenburg sentence test). *Z Audiol*, *38*, 3.

Walden, T. C., & Walden, B. E. (2005). Unilateral versus bilateral amplification for adults with impaired hearing. *Journal of the American Academy of Audiology*, *16*(8), 574–584.

Wu, Y. H., Stangl, E., Pang, C., & Zhang, X. (2014). The effect of audiovisual and binaural listening on the acceptable noise level (ANL): Establishing an ANL conceptual model. *Journal of the American Academy of Audiology*, *25*(2), 141–153.