Effects of Binaural Spatialization in Wireless Microphone Systems for Hearing Aids on Normal-Hearing and Hearing-Impaired Listeners

Gilles Courtois[1], Hervé Lissek[1], Philippe Estoppey[2], Yves Oesch[3], and Xavier Gigandet[3]

Abstract

Little is known about the perception of artificial spatial hearing by hearing-impaired subjects. The purpose of this study was to investigate how listeners with hearing disorders perceived the effect of a spatialization feature designed for wireless microphone systems. Forty listeners took part in the experiments. They were arranged in four groups: normal-hearing, moderate, severe, and profound hearing loss. Their performance in terms of speech understanding and speaker localization was assessed with diotic and binaural stimuli. The results of the speech intelligibility experiment revealed that the subjects presenting a moderate or severe hearing impairment better understood speech with the spatialization feature. Thus, it was demonstrated that the conventional diotic binaural summation operated by current wireless systems can be transformed to reproduce the spatial cues required to localize the speaker, without any loss of intelligibility. The speaker localization experiment showed that a majority of the hearing-impaired listeners had similar performance with natural and artificial spatial hearing, contrary to the normal-hearing listeners. This suggests that certain subjects with hearing impairment preserve their localization abilities with approximated generic head-related transfer functions in the frontal horizontal plane.

Keywords

spatial hearing, hearing aids, FM systems, speech intelligibility, sound localization

Introduction

Binaural hearing is a fundamental property of the human auditory system. Rather than simply replicating the information at each ear, it provides additional capabilities resulting from the analysis of the binaural sound differences. The different times of arrival of the acoustic signal (the interaural time difference), as well as the difference of sound pressure levels (SPLs; interaural level difference) between the left and right ears make possible the localization of sounds in the horizontal plane. They are combined with the monaural (spectral) cues, which occur at high frequencies and correspond to the effect of the torso, the head, and especially the pinna. These three localization cues are encapsulated in the head-related transfer function (HRTF), as described by Cheng and Wakefield (1999).

It is also well-established that binaural hearing contributes to speech intelligibility in complex and noisy conditions (Carhart, 1965). This is referred to as the cocktail party effect (Bronkhorst, 2000; Hawley, Litovsky, & Culling, 2004). The spatial release from masking denotes the intelligibility gain that is observed when a spatial separation is introduced between the targeted speech signal and the masker(s) (Dirks & Wilson,

[1]Swiss Federal Institute of Technology, Signal Processing Laboratory, Lausanne, Switzerland
[2]Acoustique Riponne, Lausanne, Switzerland
[3]Phonak Communications AG, Murten, Switzerland

Corresponding author:
Gilles Courtois, Swiss Federal Institute of Technology, Signal Processing Laboratory Acoustics, Station 11, CH-1015 Lausanne, Switzerland.
Email: gilles.courtois@epfl.ch

1969; Freyman, Helfer, McCall, & Clifton, 1999). It includes two components, which are binaural switching and binaural unmasking. The first designates the selection and focus on the ear bringing the highest signal-to-noise ratio (SNR), due to the head shadow effect (Bronkhorst & Plomp, 1988; Culling & Mansell 2013), while the second denotes the noise suppression mechanism resulting from the analysis of the noise pattern at both ears (Dillon, 2012; Gallun, Mason, & Kidd, 2005). Additionally, binaural localization helps identify the speaker and gives access to lip reading.

Hearing-impaired (HI) listeners with a bilateral hearing loss benefit less from the cocktail party effect to extract speech information from noise than normal-hearing (NH) listeners (Arbogast, Mason, & Kidd, 2005; Bronkhorst & Plomp, 1989; Noble, Byrne, & Ter-Horst, 1997), although their localization performance in the frontal horizontal plane (FHP) is generally preserved (Durlach, Thompson, & Colburn, 1981; Moore, 2007; Nordlund, 1964). Despite the significant improvement in speech understanding provided by hearing aids (HAs) compared with unaided listening, numerous studies have demonstrated that unlinked bilateral hearing devices can strongly degrade the spatial cues (Best et al., 2010; Byrne & Noble, 1998; Keidser, O'Brien, Hain, McLelland, & Yeend, 2009, Keidser et al., 2006; Noble & Byrne, 1990; Van den Bogaert, Carette, & Wouters, 2009; Van den Bogaert, Klasen, Moonen, Van Deun, & Wouters, 2006; Wiggins & Seeber, 2012). The new generation of HAs incorporates a wireless connection between the left and right devices aiming at two objectives. First, to take advantage of a network of several microphones to enhance the performance of certain signal processing algorithms, such as a beamformer or a noise canceller (Appleton & König 2014; Kreisman, Mazevski, Schum, & Sockalingam, 2010; Latzel, 2013; Timmer, 2013; Yousefian, Loizou, & Hansen, 2014). Second, to link the gain model (e.g., dynamic compression) between both devices so as to better preserve the spatial cues (Hassager, May, Wiinberg, & Dau, 2017; Picou, Aspell, & Ricketts, 2014; Sockalingam, Holmberg, Eneroth, & Shulte, 2009; Thiemann, Mller, Marquardt, Doclo, & Van de Par, 2016). However, these objectives are often conflicting and a tradeoff has to be found between SNR improvement and binaural cue preservation (Neher, Wagener, & Latzel, 2017).

A typical example of that tradeoff can be found in FM technology (in this article, the expression "FM systems" refers likewise to devices with the old analog transmission or to the most recent ones based on a digital modulation). A usual FM unit consists of a small transmitter microphone, which picks up the voice of a speaker and sends the clean speech to a radio-frequency (RF) receiver plugged into the HA of a listener, via a wireless connection. Many studies evidenced the strong intelligibility enhancements obtained with FM systems (Crandell & Smaldino, 1999; Hawkins, 1984; Lewis, Crandell, Valente, & Horn, 2004; Thibodeau, 2010, 2014), even for disorders other than hearing loss, for example, autism and hyperactivities (Schafer et al., 2013). FM systems rely on the full binaural summation of a diotic signal captured at the speaker's place. The reproduction of twice the same signal at both ears is known to increase speech intelligibility (Dillon, 2012), and FM systems take advantage of it. The counterpart is that no spatial cues are reproduced. The lack of spatial information can be partially solved in the so-called FM+M mode (as opposed to the FM-only mode), where the clean transmitted speech is mixed with the potentially noisy signal from the HA microphone, at the cost of a lower intelligibility.

The authors addressed this issue with a novel solution that preserves the high quality of the remote microphone signal, while artificially restoring the cues required for sound localization (Courtois, Marmaroli, Lissek, Oesch, & Balande, 2015a, 2015b). This process is referred to as the spatial hearing restoration feature (SHRF) in this article and is summarized in Figure 1. This figure represents a typical use case of FM systems, with two speakers wearing a remote microphone and one HI listener equipped with hearing aids and RF receivers. The algorithm first detects and localizes the current talker, with a spatial resolution of five areas in the FHP. Then, the speech from the wireless microphone is spatialized in the determined position. Thus, binaural hearing can be reintroduced without mixing the FM-transmitted speech with a noisier signal. Nevertheless, it is unknown whether the change from a full binaural summation (i.e., diotic presentation) to a partial binaural summation (i.e., spatial presentation) has an effect on speech understanding.

While the literature concerning binaural spatialization in NH listeners is abundant (see e.g., Begault, Wenzel, Lee, & Anderson, 2001; Wenzel, Arruda, Kistler, & Wightman, 1993; Wightman & Kistler 1989a, 1989b), there is a growing interest in understanding how HI listeners react to artificial spatial hearing (Best, Roverud, Mason, & Kidd, 2017; Boyd, Whitmer, Soraghan, & Akeroyd, 2012; Brungart, Cohen, Zion, & Romigh, 2017; Minnaar, 2010; Ohl, Laugesen, Buchholz, & Dau, 2010; Whitmer, Seeber, & Akeroyd, 2014). Additionally, several studies have been carried out in collaboration with HA manufacturers to investigate the attributes related to spatial perception, such as externalization, width, or distance estimation (Catic, Santurette, & Dau, 2015, Catic, Santurette, Dau, & Buchholz, 2012; Hassager, Wiinberg, & Dau, 2017; Udesen, Piechowiak, & Gran, 2015), supporting the hypothesis that their findings will lead to innovations in the hearing aid industry in a close future. The study reported in this article goes toward this
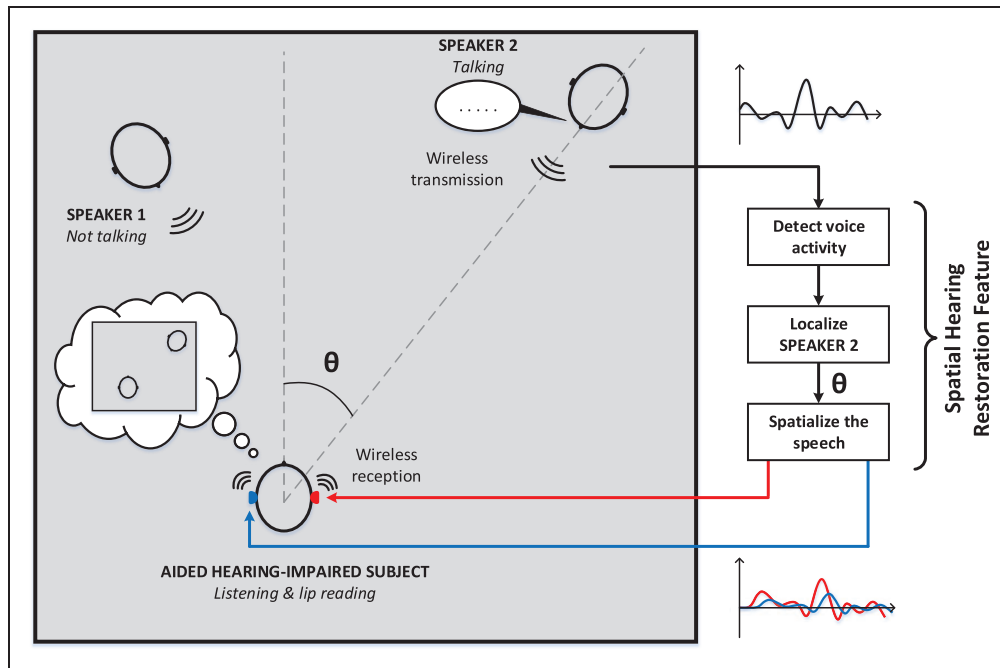
Figure 1. Principle of the SHRF that allows recovering spatial hearing in FM systems. The process combines a localization algorithm in charge of determining the position of the current talker and a binaural spatialization block to restore the corresponding spatial cues.

direction. Its first objective is to investigate the effect of reintroducing spatial cues within the FM technology. The second objective is to bring novel insights into the way HI listeners perceive artificial spatial hearing.

Methods

Participants

Forty subjects took part in this study. They were arranged in four groups:

- 10 young adult NH subjects (NH, or control group) with hearing thresholds lower than or equal to 20 dB HL at both ears (125 Hz to 8 kHz),
- 10 moderate HI subjects (HI-MOD group) with pure-tone averages (PTAs over 0.5 to 4 kHz) between 41 and 60 dB HL (all PTAs were computed at the better ear),
- 10 severe HI subjects (HI-SVR group) with PTAs between 61 and 80 dB HL,
- 10 profound HI subjects (HI-PFD group) with PTAs greater than 81 dB HL.

These categories are in agreement with the ones defined by the World Health Organization (2016). All HI subjects presented a symmetrical hearing loss that did not differ by more than 20 dB between the left and right PTAs. They were all common users of bilateral

Table 1. Statistics Related to the Four Groups of Subjects: Median Age, Range of Age, and PTAs—Mean (*SD*)—at the Better and Worse Ears.

Group	Median age	Range of age	PTA best ear (dB HL)	PTA worst ear (dB HL)
NH	21	19–24	1.6 (2.2)	3.2 (2.3)
HI-MOD	57	22–80	49.5 (4.1)	56.1 (4.4)
HI-SVR	71	20–84	67.8 (4.4)	72.3 (5.6)
HI-PFD	29	22–72	98.6 (10.4)	103.4 (9.9)

Note. PTA = pure-tone averages; NH = normal-hearing; HI-MOD = moderate hearing-impaired; HI-SVR = severe hearing-impaired; HI-PFD = profound hearing-impaired.

Phonak HAs with an available direct audio input (DAI). All participants provided written informed consent and were paid for their participation.

Table 1 reports statistics related to the four groups. The average hearing loss in the better ear was well centered in the intervals of each category, so as to avoid overlapping hearing losses between groups. Note that the difference in PTAs between the HI-MOD and the HI-SVR groups was approximately 20 dB, while the difference between the HI-SVR and HI-PFD was around 30 dB. This is because all subjects presenting a hearing loss higher than 81 dB HL were included in the HI-PFD group. The individual PTAs in this group range from 83 to 115 dB HL.
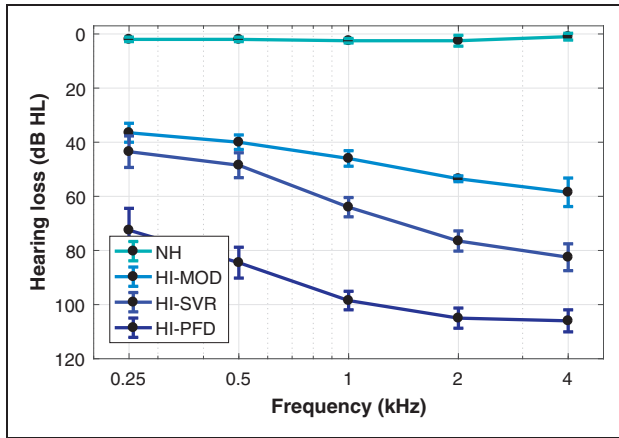
Figure 2. Average audiograms (better ear) in each category of subjects. The error bars represent the standard error.

Figure 3. The average IN/OUT characteristics of the subjects' HAs in the different groups. The error bars represent the standard error.

The average audiograms (better ear) in each group are depicted in Figure 2. Some sloping hearing losses are shown in all HI groups. The overlap between the categories is small.

Hardware

A pair of HAs (Phonak Naida IX SP) with fittings matching an audiogram with no hearing loss (linear amplification) was available for the NH subjects. Once worn, their conchae were filled with some impression material, in order to reduce the contribution of the direct sound in the ear canal. The HI subjects used their own HAs, with their usual settings of dynamic and frequency compression. For safety reasons, the feedback canceller was kept as well. All the other algorithms (adaptive processing such as reverberation canceller, noise cancellation, etc.) were switched off, and the microphone directivity was set to omnidirectional. The HI subjects kept up their usual earmolds.

Before starting the experiments, the HA of the better ear of each HI subject was submitted to a short calibration, so as to characterize its IN/OUT behavior. This procedure was in agreement with the American National Standards Institute (2003), paragraph 6.15.1 ("Input-output characteristics"), except that the signal used was either a speech sequence or a speech-shaped noise (SSN), instead of a pure-tone signal. Figure 3 shows the curves of the averaged HA dynamics in the four groups, measured in the 2 cc coupler for a speech signal, as a function of the root mean square (RMS) value of the electrical signal input. The working level (black line) is around 2 mV RMS (6 dB mV), which is the standard electrical voltage at the input of the DAI when FM systems are used. The knowledge of the IN/OUT characteristics was required to ensure accurate SNR values during the intelligibility
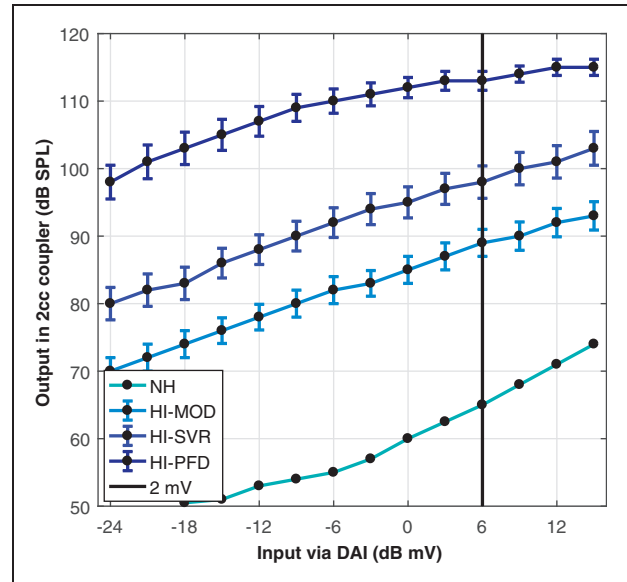
experiment. For example, in order to deliver a SNR of + 3 dB, the gain applied to the masker must be −3 dB for the NH group (linear amplification), while for the HI-PFD, a gain of −15 dB is required, due to the dynamic compression.

A MOTU 896 mk3 soundcard was used to output the audio signals from the computer. These signals were transmitted to a Denon AVR 3300 amplifier via an optical connection, so as to reduce the voltage down to 2 mV RMS. A Samson S-com plus compressor or limiter was then inserted to prevent any accidental excessive SPL. It was set to trigger at voltage levels higher than the one normally encountered in the procedure of the experiments. The attack and release times were equal to 0.3 ms and 5 s, respectively, and the ratio was turned to its "infinite" value (limiter operating mode). Finally, the signals were sent to the DAI of the HAs or to five Tannoy Reveal Active loudspeakers. The electric path stood for the FM-transmitted sound, while the loudspeakers (acoustic path) served for the experiments in the FM+M configuration. The electrical path was delayed to partly compensate for the acoustic time of flight, so that the delay between the FM-transmitted signal and the HA microphone signal was between 2 and 4 ms, as measured in the 2 cc coupler. The FM-transmitted speech was always rendered first, so that the localization was dominated by the spatialization processing (see Cranford, Andres, Piatz, & Reissig, 1993, for a review about the precedence effect). In real applications of FM systems, it might happen that the sound picked up by the HA microphone arrives the first, in case of short speaker-to-listener distances, but the FM-transmitted

speech is always reproduced louder (typically 10 to more than 20 dB higher, the so-called FM-advantage; Thibodeau, 2014). This is to guarantee a high intelligibility even in very noisy situations. The consequence is that the speech localization is usually governed by the FM-transmitted speech.

Stimuli

The SHRF splits the FHP into five so-called *spatial sectors*. This is the actual spatial resolution of the localization and spatialization processing. One sector stands for the central positions of the speaker ($-20°$ to $20°$), while two lateral sectors are available on the left and right sides. These are the "intermediate" sectors that cover the speaker positions from $20°$ to $40°$ (resp. $-20°$ to $-40°$ on the left) and the "extreme" sectors corresponding to the speaker positions between $40°$ and $90°$ (resp. $-40°$ to $-90°$). The FM-transmitted speech is spatialized at $0°$ in the central sector, at $\pm 30°$ in the intermediate sectors, and at $\pm 65°$ in the extreme sectors, using 10th-order minimum-phase finite impulse response filters. Figure 4 displays the gain of those filters as a function of the frequency (blue and red solid lines), for a sound spatialized at $-65°$. Their frequency–gain responses are compared with the original HRTFs that were used to design the filters at this azimuth (dashed lines). These HRTFs were taken from the CIPIC database by Algazi, Duda, Thompson, and Avendano (2001), using the Subject 21 (KEMAR with large pinna) at $0°$ elevation. The bandwidth has been reduced to 8 kHz (usual bandwidth of FM systems). To avoid excessive binaural gain differences that would reduce the benefits coming from the full binaural summation in conventional FM systems, some amplitude limitation was included in the spatial filters, so that the interaural gain difference reached 20 dB at most. This processing has been described and validated by Courtois et al. (2016) on 38 NH listeners.

Two French databases of speech content were used: the HINT database by the Collège National d'Audioprothèse (2006) for the intelligibility experiment and the SUS database by Raake and Katz (2006) for the localization experiment. The first consists of meaningful and phonetically balanced sentences with four to seven words. The postulate was to resort to meaningful material in order to get closer to the speech understanding in real life, where listeners can exploit their cognitive abilities for guessing possible missing words. On the contrary, the second database is composed of semantically unpredictable (i.e., meaningless) sentences. This material was preferred to ensure that the subjects did not concentrate on the content, but rather on the direction of arrival (DOA) of the voice. The speech stimuli were spatialized with the spatial filters included in the SHRF. Prior to the experiments, the long-term RMS value of a SSN had
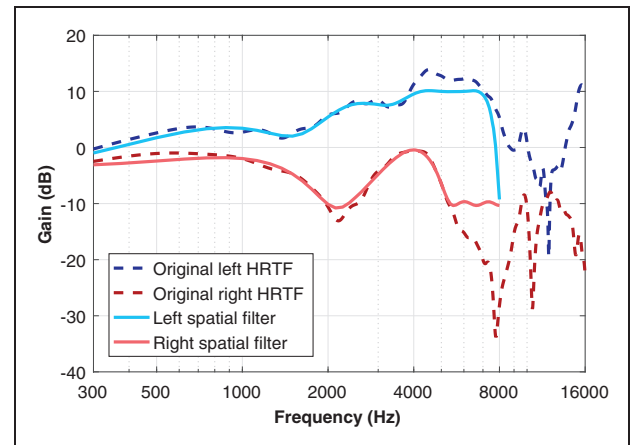

Figure 4. The spatial filters (solid lines) used for a sound image at $-65°$ compared with the original HRTFs (dashed lines).

been measured in the 2-cc coupler when it was either diotic or spatialized and played via the DAI. The levels had been adjusted until the respective computed loudnesses were the same, according to the model of Glasberg and Moore (2002). The other spatialized directions were supposed to yield the same binaural SPL as the one at $0°$ (Begault & Erbe, 1994). This procedure ensured an equal loudness of the diotic and spatialized speech.

The masker was a mixture of five uncorrelated SSNs having the same long-term spectrum as the HINT corpus. Each of these five noises was spatialized in one of the five sectors. Hence, the spatialized speech signal was always colocated with one of the five noise signals. This was to ensure a limited contribution of the spatial release from masking in the intelligibility assessment, so that the potential effect of the processing on speech understanding would be primarily attributed to the change from a diotic to a dichotic presentation.

For the NH group, the listening level was set to 65 dB SPL. The gain of the HAs had thus been adjusted so that an electrical input at 2 mV RMS corresponded to 65 dB SPL in the ear canal. For the HI subjects, the same input voltage was used, and the listeners' personal-fitted gain delivered the speech at a comfortable acoustic level. A variation of ± 8 dB around the 2 mV voltage was possible if requested by the subject. When the FM+M configuration was considered, the stimuli were simultaneously played via the DAI and through the loudspeakers. Both paths equally contributed to obtain 65 dB SPL at the output of the HA.

Procedure

Intelligibility experiment. The goal of this experiment was to evaluate the impact of the SHRF on the understanding of speech. More precisely, it was intended to study the

effect of the change from a diotic (i.e., full binaural summation) to a spatialized (partial binaural summation) sound reproduction. It was performed in two different configurations, corresponding to the FM-only and FM+M modes, respectively. In the first configuration (FM-only), some stimuli were randomly spatialized in one of the five spatial sectors, while some others were left unprocessed (i.e., diotic). In the second configuration (FM+M), some stimuli were simultaneously spatialized in one of the five locations (FM path) and played in the corresponding loudspeaker, to be captured by the HA microphones (M path). When the stimuli were rendered diotic, a random loudspeaker played the sound at the same time.

The listeners sat in the center of the room (background noise < 25 dBA, $RT_{60} = 0.17$ s, volume = $125\,m^3$). They were asked to look straight ahead, and their head was immobilized by a chin rest. For both configurations, the sentences were played at three different SNRs. The masking noise was always input via the DAI, and all SNRs were adjusted by varying the noise level, while keeping the speech level fixed. The same masking noise (i.e., mixture of the five spatialized SSNs) was used for both the diotic and spatialized sentences in the two configurations. There were two sentences per direction and three diotic ones, giving a total of 39 sentences (3 SNRs ×[3 diotic + 2 × 5 directional conditions]). The processing of every sentence (i.e., diotic or spatialized locations) was randomized within and across subjects. After each sentence, the listeners were asked to repeat what they understood, and the correct words were collected to derive the speech recognition score (SRS), that is, the percentage of correctly repeated words in the sentence. The NH and HI subjects were not tested with the same SNRs. The NH group experienced SNRs of −10 dB, −13 dB, and −16 dB. For each HI listener, the procedure suggested by Lewis et al. (2004) was adopted and adapted to the experiment. The examiner fixed a starting SNR for the 13 first sentences, that was equal to −6 dB for the HI-MOD listeners, −3 dB for the HI-SVR listeners, and 0 dB for the HI-PFD listeners. Then, depending on the results, two other SNRs were driven by 3-dB steps, such that:

- SNR HIGH yielded the best intelligibility score,
- SNR MID yielded an intermediate intelligibility score,
- SNR LOW yielded the worst intelligibility score.

For example, if a moderate HI listeners reached a SRS of about 50% after the 13 first sentences (SNR at −6 dB), the two tested SNRs were −3 dB and −9 dB. In the case where a listener provided an intelligibility score close to 0% at the first SNR, he was then submitted to SNRs at −3 dB and 0 dB. For moderate listeners presenting a high-intelligibility performance with the initial SNR (i.e., SRS near 100%), they experienced following SNRs at −9 dB and −12 dB. This procedure was chosen because it avoids encountering undesired floor or ceiling effects (i.e., 0% or 100% intelligibility scores). In the HI-PFD group, only six profound listeners managed to pass the experiment, while the four others could not understand any word, even when no noise was added.

Each configuration started with a training period of six test sentences (one diotic + one in each sector), such that the listeners could get used to the procedure and hear the various spatial conditions once. The subjects were not aware of the actual start of the experiment after this training time.

Localization experiment. The objective of the localization experiment was to evaluate the effect of the binaural spatialization, as introduced by the SHRF, on the localization performance of NH and HI listeners. This was done in four configurations:

- Configuration 1 (Unaided): The subjects were tested without their HAs. The acoustic level was adjusted for each HI listener to be sufficiently loud,
- Configuration 2 (Aided): The subjects were tested with their HAs and usual fittings,
- Configuration 3 (FM-only): The subjects were tested when the spatialization was played back via the DAI only,
- Configuration 4 (FM+M): The subjects were tested when the spatialization was rendered via the DAI, and simultaneously through the corresponding loudspeaker.

The listeners could not see the loudspeakers, which were hidden by a black curtain, as shown on Figure 5. Nine numbers from −4 (left) to 4 (right) were displayed on the curtain, corresponding to azimuths at $0°$, $\pm15°$, $\pm30°$, $\pm45°$, and $\pm65°$. This procedure was identical in the four configurations.

In this experiment, three sentences were played in each direction, resulting in a total of 15 sentences per configuration. The processing of every sentence (i.e., spatialized locations) was randomized within and across subjects. After each sentence, the listeners were asked to indicate the perceived location of the sound source by reporting the number corresponding to the incidence direction (see Figure 5). Then, the localization error was computed as the difference between the number reported by the listener and the number of the actual loudspeaker emitting the sound. Note that the localization error was not computed in degree due to the coarse resolution of the SHRF, and the fact that the sectors do not present the same angular span. The listeners were made aware that all the available azimuths may not be played.
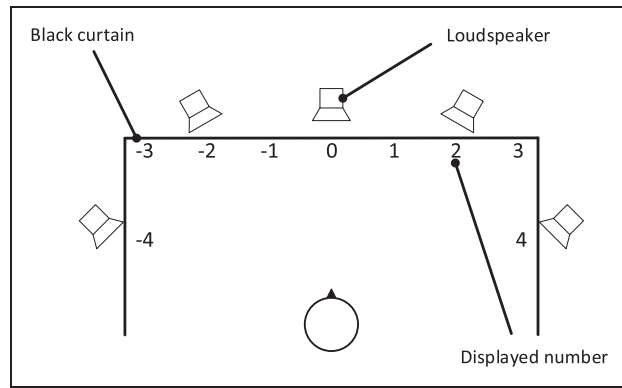
Figure 5. Setup of the localization experiment, showing the loudspeakers hidden by the curtain and the visible numbers.

Table 2. HIGH, MID, and LOW Average SNRs (in dB) Experienced by the Listeners in the Four Groups.

Group	HIGH	MID	LOW
NH	−10	−13	−16
HI-MOD	−3.9 (3.2)	−7 (3.4)	−9.7 (3.5)
HI-SVR	−1.3 (2.2)	−4.3 (2.1)	−7.3 (2.2)
HI-PFD	5 (4.5)	3 (8.8)	1 (4.3)

Note. Standard deviations are given into brackets. PTA = pure-tone averages; NH = normal-hearing; HI-MOD = moderate hearing-impaired; HI-SVR = severe hearing-impaired; HI-PFD = profound hearing-impaired; SNR = signal-to-noise ratio.

They could also answer that they perceived the sound from none of those directions (e.g., above, behind, etc.).

All configurations started with a training period of five test sentences in each spatialized direction, so that listeners could get used to the procedure and hear the various spatial conditions once. The subjects were not aware of the actual start of the experiment after this training time. A roving level of ± 3 dB among stimuli was implemented throughout the experiment. This was to minimize the risk that the listeners rely on potential differences of loudness among locations to infer the position of the sound source, see, for example, Noble et al. (1997), Keidser et al. (2006), and Majdak, Walder, and Laback (2013). All HI subjects managed to perform this test.

Results

Intelligibility and SNR

Table 2 reports the average SNRs experienced by the listeners in the four groups in the HIGH, MID, and LOW intelligibility conditions. The performance of the NH listeners was tested at fixed SNRs, while individualized SNRs were used for the HI subjects. While the average step between the three conditions is around 3 dB for the HI-MOD and HI-SVR group, similarly to the NH group, this step is reduced to 2 dB in the HI-PFD group. This is because, the profound HI listeners presented a narrow SNR interval between a satisfying and a null speech understanding, limiting the SNR range that could be tested.

Figure 6 shows the SRS in the four groups for the three tested SNRs (HIGH in blue, MID in orange, and LOW in red) in the FM-only (a) and in the FM+M (b) configurations, averaged over all rendering types (diotic and spatialized). The absolute scores must not be compared between groups because the SNRs were specific to each HI listener. The outcomes from a two-way repeated measures analysis of variance (ANOVA) investigating the influence of the SNR and configuration are reported in Table 3. The statistical analysis revealed that there was a significant effect of the SNR on the speech understanding. Specifically, the speech intelligibility decreased when the SNR diminished for all groups, despite the strong intragroup variations that could be observed in the HI groups, most probably due to the procedure of individualized SNRs. In the NH, HI-MOD, and HI-SVR groups, three different distributions of SRS can be distinguished, with limited overlap. The test failed to show any statistical influence of the configuration. There was no interaction effect between both factors in all groups.

Intelligibility and Rendering

Figure 7 displays the SRS distributions obtained with the usual diotic rendering (yellow) and the spatial processing from the SHRF (green), in the FM-only (a) and FM+M (b) configurations. The SRS is presented in each group and averaged over all SNRs. In the FM-only configuration, the graph suggests an improvement of the intelligibility with the SHRF in the NH, HI-MOD, and HI-SVR groups. In the FM+M configuration, one can suspect an enhancement of the speech understanding for the moderate HI subjects. On the contrary, a diminution of the intelligibility might occur in the HI-PFD group when the SHRF is applied.

Tables 4 and 5 report the results from the two-way repeated measures ANOVA investigating the influence of the rendering and the interaction effect between the rendering and SNR on the speech intelligibility in the FM-only and FM+M configurations, respectively. The goal was to test the alternative hypothesis that the spatialization feature does improve speech intelligibility, against the null hypothesis assuming that it has no influence. If the alternative hypothesis was accepted, it was also desired to know whether the benefit coming from
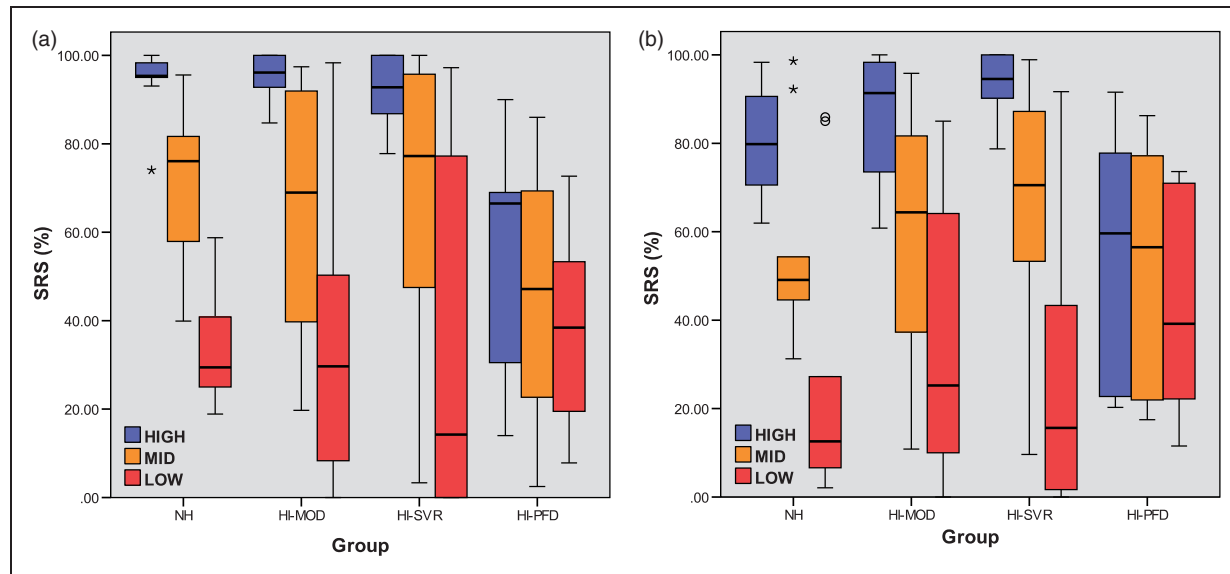
Figure 6. Distribution of the SRS as a function of the SNR for the different groups in the FM-only (a) and FM+M (b) configurations. The SRS are averaged over all rendering types (diotic and spatialized). The dark line in the middle of the boxes represents the median, while the bottom and top lines correspond to the 25th and 75th percentiles, respectively. The T-bars are the whiskers (correspond to the min and max values within 1.5 times the interquartile range). The points are outliers and the stars indicate extreme outliers (values more than three times the height of the box).

Table 3. Results of a Two-Way Repeated Measures ANOVA, Showing the Effect of the SNR and Configuration on the Speech Intelligibility for the Four Groups.

Group	Factor	df 1	df 2	F	p
NH	SNR	2	18	70.30	**<.001**
	Config	1	9	2.357	.159
HI-MOD	SNR	2	18	28.62	**<.001**
	Config	1	9	1.305	.283
HI-SVR	SNR	2	18	27.38	**<.001**
	Config	1	9	0.432	.528
HI-PFD	SNR	2	10	4.516	**.04**
	Config	1	5	1.441	.284

Note. The significant effects are given in bold ($\alpha = .05$). PTA = pure-tone averages; NH = normal-hearing; HI-MOD = moderate hearing-impaired; HI-SVR = severe hearing-impaired; HI-PFD = profound hearing-impaired; SNR = signal-to-noise ratio.

the SHRF appeared in certain SNRs specifically. In the FM-only configuration, the analysis confirmed a significant enhancement of the speech understanding performance for the moderate and severe HI listeners, while this was not observed in the NH and HI-PFD groups. In the FM+M configuration, a significant effect of the SHRF on the intelligibility was present in the HI-MOD group only. No interaction effect between the rendering and the masker level was found in any case.

Intelligibility and DOA

The influence of the DOA of the speaker's voice on the SRS was investigated. A sequence of one-way repeated measures ANOVAs was performed to look for significant effects of the spatial sectors in the results of the four groups. In the FM-only configuration, a statistical influence was found for the HI-SVR group, $F(5, 45) = 2.998, p = .020)$. In the FM+M configuration, there was a statistical effect with the moderate HI listeners, $F(5, 45) = 2.579, p = .039$. However, no effect remained after having applying a Bonferroni correction for multiple comparisons.

Localization and Configuration

The localization error in the four configurations is presented in Figure 8. Considering the NH group, the figure suggests that there exist some significant differences between the configurations. This was confirmed by a one-way repeated measures ANOVA, $F(3, 27) = 3.837$, $p < .05$. A one-tailed Bonferroni post hoc test indicated that there was a degradation of the localization performance between the unaided and FM-only configurations ($p = .036$). In the three HI groups, no significant difference was found between the configurations: HI-MOD: $F(3, 27) = 0.960, p = .426$; HI-SVR: $F(3, 12.505) = 0.951, p = .380$; HI-PFD: $F(3, 27) = 0.510, p = .679$. A Greehouse–Geisser correction was applied for the HI-SVR group due to sphericity assumption violation.
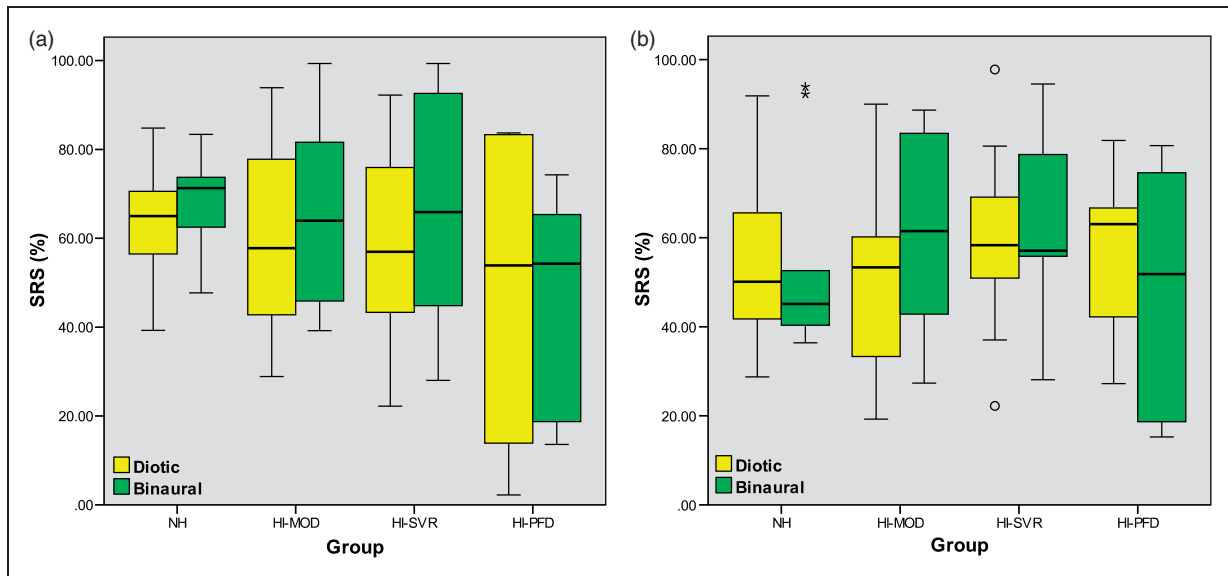
Figure 7. Distribution of the SRS in the four groups, with the diotic (yellow) and binaural (green) renderings. (a): FM-only configuration; (b): FM+M configuration (B). The SRS are averaged over all SNRs.

Table 4. Results of a Two-Way Repeated Measures ANOVA, Showing the Effect of the Rendering and the Interaction Between SNR and Rendering on the Speech Intelligibility for the Four Groups in the FM-Only Configuration.

Group	Factor	df 1	df 2	F	p
NH	Rend.	1	9	0.750	.409
	SNR × Rend.	2	18	2.927	.079
HI-MOD	Rend.	1	9	10.94	**.009**
	SNR × Rend.	2	18	0.709	.506
HI-SVR	Rend.	1	9	19.96	**.002**
	SNR × Rend.	2	18	1.319	.292
HI-PFD	Rend.	1	5	0.149	.715
	SNR × Rend.	2	10	1.074	.378

Note. The significant effects are given in bold ($\alpha = .05$). PTA = pure-tone averages; NH = normal-hearing; HI-MOD = moderate hearing-impaired; HI-SVR = severe hearing-impaired; HI-PFD = profound hearing-impaired; SNR = signal-to-noise ratio; Rend. = rendering.

Table 5. Results of a Two-Way Repeated Measures ANOVA, Showing the Effect of the Rendering and the Interaction Between SNR and Rendering on the Speech Intelligibility for the Four Groups in the FM+M Configuration.

Group	Factor	df 1	df 2	F	p
NH	Rend.	1	9	0.305	.207
	SNR×Rend.	2	18	0.718	.064
HI-MOD	Rend.	1	9	9.298	**.014**
	SNR×Rend.	2	18	3.009	.075
HI-SVR	Rend.	1	9	0.600	.428
	SNR×Rend.	2	18	0.084	.920
HI-PFD	Rend.	1	5	1.304	.305
	SNR×Rend.	2	10	0.343	.718

Note. The significant effects are given in bold ($\alpha = .05$). PTA = pure-tone averages; NH = normal-hearing; HI-MOD = moderate hearing-impaired; HI-SVR = severe hearing-impaired; HI-PFD = profound hearing-impaired; SNR = signal-to-noise ratio; Rend. = rendering.

Localization and DOA

Figure 9 depicts the influence of the DOA of the speaker's voice on the localization error. The central (CTR, 0°), intermediate (INT, ±30°), and extreme (EXT, ±65°) sectors are considered as spatial locations. Some significant differences between the three types of sectors can be suspected. In particular, it seems that the intermediate locations led to worse performance compared with the central and extreme sectors, apart from the HI-PFD group. A sequence of one-way repeated measures

ANOVAs with Bonferroni correction for multiple comparisons confirmed the significant effect of the sector in the NH, $F(2, 18) = 10.76, p < .05$, HI-MOD, $F(2, 18) = 10.87, p < .05$, and HI-SVR, $F(2, 18) = 6.368, p < .05$, groups. Specifically, the localization performance in the intermediate sectors was significantly worse than in the extreme sectors in the NH group ($p = .026$), and the performance of the moderate and severe HI listeners was significantly worse in the intermediate sectors than in the central sector (HI-MOD: $p = .005$; HI-SVR: $p = .017$). The proportion of answers reporting a
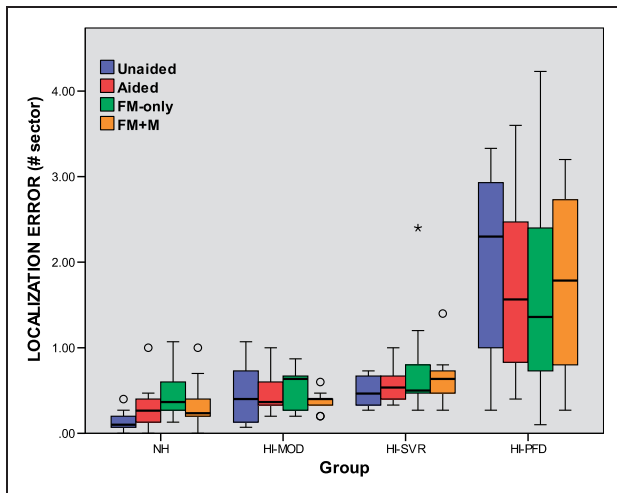
Figure 8. Distribution of the localization error in different configurations for the four groups.
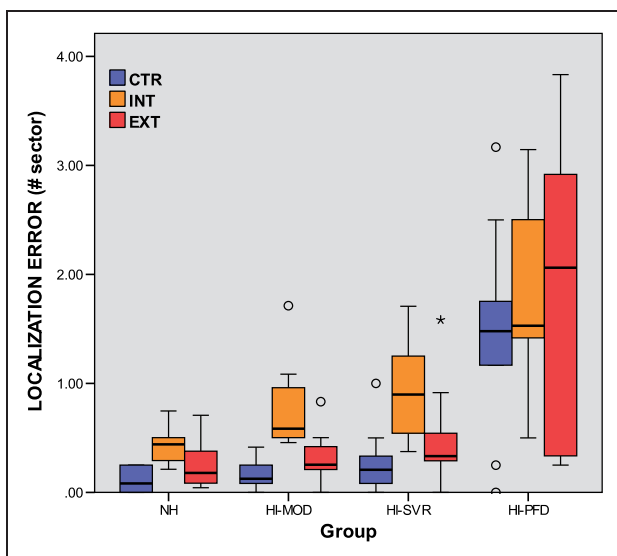

Figure 9. Distribution of the localization error in the central, intermediate, and extreme sectors for the four groups.

perception of the source from above or behind represented 2.8% over the total answers and were mostly encountered in the HI-PFD group.

Localization and Group

Contrary to the intelligibility test, it is possible to compare the localization performance between groups, as shown in Figure 8. To this end, several one-way between-subjects ANOVAs were conducted. First, the HI-PFD group was not considered in the statistical analysis, as it yielded a violation of the assumption of variance homogeneity. The null hypothesis stood that there

was no significant difference of localization performance between the three other groups, while the alternative hypothesis claimed that there were significant variations with the degree of hearing loss. Statistical effects were found in the unaided, $F(2, 27) = 7.453, p = .002$, aided, $F(2, 27) = 2.729, p = .037$, and FM+M, $F(2, 27) = 4.664, p = .009$, configurations. Tukey's one-tailed post hoc tests showed a significant increase of the localization error between the NH and HI-MOD groups ($p = .006$) and between the NH and HI-SVR groups ($p = .002$) in the unaided configuration. Then, only statistical differences between the NH and HI-SVR groups were observed in the aided ($p = .031$) and FM+M ($p = .013$) configurations. Comparing the two HI groups, a significant degradation of the performance arose between the moderate and severe HI subjects in the FM+M configuration ($p = .024$). No significant differences were found between the three groups in the FM-only configuration.

The analysis of the results in the HI-PFD required to resort to another procedure because the data did not fulfill the assumption of variance homogeneity. A one-way between subjects ANOVA including a Brown–Forsythe correction for unequal variances showed a significant degradation of the localization performance between the severe and profound HI subjects, in all configurations: Unaided: $F(1, 8.307) = 14.457, p = .003$; Aided: $F(1, 9.707) = 11.257, p = .004$; FM-only: $F(1, 13.205) = 3.9, p = .035$; FM+M: $F(1, 10.267) = 7.830, p = .009$.

Age and PTA

Although there was a high variability of age between subjects in each group, as described in Table 1, no significant correlation between age and performance was found in the intelligibility experiment for either configuration (HI-MOD: $r = -.098$, $N = 10$, $p = .787$; HI-SVR: $r = -.160$, $N = 10$, $p = .658$; HI-PFD: $r = .128$, $N = 6$, $p = .810$) or in the localization experiment for any configuration (HI-MOD: $r = .030$, $N = 10$, $p = .935$; HI-SVR: $r = .209$, $N = 10$, $p = .562$; HI-PFD: $r = .227$, $N = 10$, $p = .528$). Nevertheless, a significant correlation was found between the PTAs at the better ear and the localization performance in the four configurations ($r = .683$, $N = 40$, $p < .001$), as could be expected.

Discussion

Intelligibility and SNR

The procedure adapted from Lewis et al. (2004) has yielded a powerful way of conducting intelligibility experiments, by finding the adequate SNRs for every HI subject, while avoiding some undesirable floor and ceiling effects. Lower SNRs produced worse speech

understanding performance in each group, as expected. In the FM-only configuration, the general distribution of the SRS with the various SNRs follows the same trend in the NH, HI-MOD, and HI-SVR groups, with median scores around 95%, 70%, and 30% for the HIGH, MID, and LOW SNRs, respectively. This is consistent with the fact that these three groups experienced the same steps of 3 dB between each SNR.

The variations of the performance within the HI-MOD and HI-SVR groups appear to be much higher than in the NH group. The procedure of customized SNRs across the HI subjects is hypothesized as the main reason for that (see Table 2). In fact, the SNRs experienced by the subjects in the HI-MOD group varied from 0 to −9 dB for the HIGH SNR, from −3 to −13 dB for the MID SNR, and from −6 to −15 dB for the LOW SNR, depending on their individual performance. In the HI-SVR group, it varied from 3 to −3 dB (HIGH), from 0 to −6 dB (MID), and from −3 to −9 dB (LOW). Additionally, it is well known that performance between HI subjects that have similar PTAs can dramatically differ, because PTA is not always well correlated with speech perception (Smoorenburg, 1992).

The six subjects presenting a profound hearing impairment were more sensitive to the changes of the masker level than the other HI subjects. Indeed, a variation of the SNR by 3 dB could make their speech intelligibility performance falling from 100% to 0%, somehow reducing the efficiency of the individualized-SNR procedure. This is illustrated in Figure 6, where the distributions of the SRSs in the HI-PFD group are strongly overlapping. It has been evidenced by Duquesnoy and Plomp (1983) that the detrimental effect of interfering noise on speech understanding becomes stronger and faster as the PTA of HI subjects rises. The results obtained in this intelligibility experiment are in agreement with their conclusions.

Intelligibility and Rendering

One of the main results of this study is that the SHRF significantly improved speech intelligibility of the subjects presenting a moderate hearing impairment in both FM-only and FM+M configurations by an average amount of 9%. In more detail, 9 subjects over 10 experienced an increase of the intelligibility between 2% and 17%, while only one subject presented a marginal loss of 0.9%. In the FM-only configuration, all the severe HI subjects improved their speech understanding performance with the SHRF, from 0.5% to 18%. Although there was no overall significant effect of the SHRF in the FM+M configuration, the individual results showed that only half of the subjects still benefited from the spatial processing (range +1% to +31%), while the five others experienced a degradation of their speech

intelligibility (−2% to −13%). No general conclusion could be drawn from the NH and HI-PFD groups. In the latter, three subjects over six (FM-only mode) and two subjects over six (FM+M mode) experienced a gain in speech understanding with the SHRF.

The reported results show that the restoration of spatial hearing as achieved by the SHRF enhances the speech understanding performance of a strong majority of the subjects presenting a moderate or severe hearing loss. This cannot be attributed to the introduction of a spatial separation between the targeted speaker and the masking noise, since the masker was spatially distributed all over the FHP. Furthermore, the procedure of loudness compensation that was performed between the diotic and spatialized rendering ensured that this outcome was not the consequence of higher SNRs artificially introduced by the SHRF. It rather means that the full binaural summation, which is achieved by conventional FM systems, can be slightly modified to incorporate the binaural cues corresponding to the position of the speaker, without lowering the speech intelligibility. Here, the passing from a diotic to a binaural rendering was operated cautiously, with the use of amplitude-limited HRTFs, so that the gain difference applied between both HAs never went over 20 dB (see Courtois et al., 2016, for more detail).

The second conclusion that can be drawn from this experiment is that the improvements offered by the SHRF may be reduced when the processed speech is mixed with the acoustic signals captured by the HA microphone. It is well established that the FM+M configuration tends to reduce the speech perception performance that can be obtained when the FM-transmitted voice is played alone (FM-only; Thibodeau, 2010, 2014), because the ''M path'' adds part of the acoustic noise present in the environment. However, the SNR was kept similar in both configurations of this intelligibility experiment, and the masker was played through the DAI only. One can hypothesize that this observation might be rather due to the interaction between the artificial spatialization coming from the SHRF and the natural spatial hearing provided by the HA microphones. Depending on the resemblance between the HRTFs of the subjects and the ones used to design the spatial filters in the SHRF, some conflicting binaural information could lead to deteriorated speech understanding. This may also explain the strong disparity of performance that was observed between subjects inside each group.

Localization and Configuration

The experiment showed that the localization performance of the NH group was degraded in the FM-only configuration (spatialization based on generic HRTFs with the SHRF). In more detail, all the NH listeners

experienced an increase of their localization error by an average factor of 3.5, compared with the unaided configuration. This does not match the results reported by Wenzel et al. (1993) and Begault et al. (2001), who evidenced that NH listeners do not need to hear with their own HRTFs to preserve their localization abilities in the FHP. Drullman and Bronkhorst (2000) observed no difference on the localization performance of NH listeners in the FHP, despite a reduced bandwidth at 4 kHz. Similar outcomes were mentioned by Majdak et al. (2013) for the entire 3D plane, with band-limited HRTFs (8.5 kHz). However, the spatial filters available in the SHRF are only approximations of the original generic HRTFs, due to the 10th-order filter design and the gain limitation (Figure 4). It is likely that such alterations in the quality of the spatial rendering were sufficient to lower the performance of the NH listeners. The use of HAs (instead of, e.g., headphones) to present artificial spatialization may be responsible of this deterioration as well, since headphones provide a better sound quality and high fidelity. Finally, it must be noted that the spatial resolution in the present investigation was quite a bit coarser than in the aforementioned studies. It is likely that a long-term training with these spatial filters, so that subjects learn to combine vision and hearing stimuli, would result in better performance in localization, as shown by, for example, Mendonca et al. (2012) and Majdak et al. (2013). Thus, this result does not exclude the use of the SHRF on NH subjects using FM systems.

The statistical analysis did not allow to draw general conclusions on the effect of the SHRF on the localization performance in the HI groups. Looking at the individual results, it appeared that 15 HI subjects over 30 experienced an improvement of their localization abilities with the processing achieved by the SHRF, compared with the aided condition (i.e., no spatialization). They were five in the HI-MOD group, three in the HI-SVR group, and seven in the HI-PFD group. Interestingly, the reintroduction of the HA microphone signals in the FM+M configuration made the localization error higher in 10 of them again. Additionally, the performance of five subjects was similar between the natural and artificial spatial rendering. Two hypotheses can be established from these results. First, it seems that a significant part of the HI subjects did not need their own HRTF to localize sounds in the FHP. Second, the SHRF might provide more precise localization cues than the usual HA sound reproduction, especially for those with a high degree of hearing loss. However, in 10 over the 30 tested HI subjects, the localization error rose when the SHRF was activated. Like in the intelligibility experiment, this might be associated with subjects whose HRTFs are much different than the generic ones used to design the spatial filters. The continuous training

that would naturally occur if the SHRF would be used with the available visual cue should provide enhancements of the localization performance.

Brungart et al. (2017) were one of the first to evaluate the effect of artificial spatialization on HI listeners. They demonstrated that both NH and HI subjects performed better with the natural rather than the artificial spatial hearing in the entire horizontal place. However, it is difficult to compare their results with the ones reported here, since the protocol and the way of assessing the performance were substantially different. Indeed, Brungart et al. (2017) tested their subjects without hearing aids, by comparing their unaided localization performance with the one obtained by virtual playback through headphones, and their head movements were tracked to update the spatialization processing in real time. When looking at the detailed results, it is shown that a great amount of the localization error was due to front/back and back/front reversals, which were not investigated in the present study. It is therefore inappropriate to state that the observations reported in the current experiment contradict the ones from Brungart et al. (2017). The use of artificial spatialization in HI listeners is still at its early stages, and much more research is required to draw general conclusions.

Localization and DOA

The analysis of the effect of the DOA on the results revealed some differences of performance between the spatial sectors. The localization error was significantly worse in the intermediate sectors than in the central or extreme ones in several conditions. Many listeners reported that it was difficult to make a choice between the directions 1, 2, and 3 (resp. −1, −2, and −3 on the left side). It was expected that the localization performance would decrease as the source moved from the frontal to the lateral azimuths, as a consequence of the spatial resolution of the human auditory system (Blauert, 1997). Yet, the results showed that the accuracy was better in the extreme sectors than in the intermediate ones. This was most probably due to a bias in the protocol, because the listeners could not give a perceived position beyond −4 and 4. With the headrest, the subjects were barely able to see the number ±4. Hence, no additional answers (e.g., ±5) could have been added without enabling head motions. The effect of this bias must be tempered by the fact that the listeners were allowed to report a perceived DOA in none of the available positions, and this type of answer remained extremely rare (2.8% of the total number of reported locations); 66% of those responses occurred in the HI-PFD group, where no significant difference was found between the intermediate and extreme sectors.

Localization and Group

The localization experiment showed that the NH subjects performed significantly better than the HI subjects in the unaided configuration, despite the SPL compensation. More precisely, the average localization error of the NH group was multiplied by 3.4 in the HI-MOD group, by 3.7 in the HI-SVR group, and by 15 in the HI-PFD group. When subjects were equipped with the HAs, this discrepancy diminishes, and no significant difference in the localization performance remained between the NH and the HI-MOD groups. However, it cannot be concluded from these results that the use of HAs restored the localization abilities of the moderate HI subjects to some "normal" performance, because a majority of NH subjects experienced more difficulties to localize the speaker when wearing the HAs. This can be attributed to the fact that the NH subjects had no time to acclimatize to the unusual listening condition they were encountering in this configuration, that is, with conchae closed, pinna filtering shortcut by the microphone location, and sound played back through HAs. When looking at the individual results, it appeared that half of the 30 HI performed better without their fitted HAs than with.

Comparing the HI groups, it was shown that the moderate and severe HI subjects presented close performance, except in the FM+M configuration. In the HI-PFD group, the localization error considerably rose. Wiggins and Seeber (2012) evidenced that the auditory system is still capable of adapting and preserving correct localization performance to a certain extent, especially for broadband stimuli. This might explain why the subjects in the HI-MOD and HI-SVR groups localized sound with a similar accuracy. However, this adaptation seems to be insufficient to maintain satisfying localization performance when the hearing loss reaches high degrees.

Study Limitations

One objective of this study was to end up with conclusions that would be somehow generalizable to various kinds of HI subjects. The main subject-dependent factor that had been retained in the protocol was the degree of the hearing loss, but the dispersion of the performance inside each category, as well as the limited number of effects with statistical significance that were found, suggest that other factors should be considered in further investigations. This may include the age of the subjects, the origin of their hearing loss (congenital/presbysusis), the degree of symmetry between both ears, and so forth. It has been shown that the processing of speech cues, such as the temporal fine structure analysis (Hopkins & Moore, 2011), the neural representation (Tremblay, Piskosz, & Souza, 2003), or the binaural

interactions (Neher, 2017), gets poorer with age, even though it is difficult to clearly distinguish between the effects of aging and age-related hearing loss on speech understanding (Divenyi, Stark, & Haupt, 2005; Plomp & Mimpen, 1979). Due to a disordered processing of the interaural time difference, interaural level difference, and precedence effect, the localization of sound events is also slightly, but significantly, worse in the elderly (Abel, Gigure, Consoli, & Papsin, 2000; Cranford et al., 1993; Dobreva, O'Neill, & Paige, 2011). Nevertheless, no effect of age was found in the localization experiment of the current study, and the worsening of the performance was primarily attributed to hearing loss.

Another limitation of this study was the choice to keep the dynamic and frequency compressions as they were fitted for each HI listener. That prevented from clarifying their respective effect on the observed intelligibility and localization performance. Although these processing are known to distort the reproduction of the binaural and spectral cues (see e.g., Keidser et al., 2006; Van den Bogaert et al., 2009; Wiggins & Seeber, 2012), they also bring a proved advantage for speech understanding in aided subjects (Bohnert, Nyffeler, & Keilmann, 2010; McCreery et al., 2014; Moore, Peters, & Stone, 1999). Two motivations supported their preservation. First, the tested HI subjects were accustomed to hearing with their own fittings and may have been disturbed if those processing were switched off. Second, the SHRF would always be followed by frequency and dynamic compressions, as achieved by the HA. Therefore, the reported protocol got closer to real-life listening conditions.

The lack of realism of certain listening scenarios can be seen as a drawback of this study. Although absent in nature, the SSN was chosen since it is frequently used in the literature (see e.g., Culling & Mansell, 2013; Drennan, Gatehouse, Howell, Van Tasell, & Lund, 2005; Duquesnoy & Plomp, 1980; George, Festen, & Houtgast, 2006) and known as the most difficult masker to cope with (Lewis et al., 2004). A more realistic interfering noise could have been an isotropic babble noise. The presentation of the masker via the DAI rather than through loudspeakers also appears to be unrealistic and was motivated by two reasons. First, it avoided the occurrence of any feedback, even at high noise levels. Second, the precision of the desired SNR was quite a bit better through the DAI, thanks to the prior measurements of the IN/OUT characteristics of the HAs of each patient. Finally, the use of static spatialization presentations contributed to an unnatural sound reproduction as well. The use of a head tracker, combined with a dynamic spatialization processing would have provided more realistic listening scenarios. A significant number of severe and profound HI subjects insisted upon the fact that lip reading constituted

a prominent help for understanding speech in their daily life, while they could not resort to this cue in the intelligibility experiment. Audiovisual stimuli including a movie showing the face and lips of the speaker should be considered as a complement for voice in further researches, as done by Macleod and Summerfield (1990) and Beskow et al. (1997). An evaluation of the time required by subjects to identify and have access to lip reading should be performed as well. In a complex auditory scene including background noise or several potential talkers, it is expected that this duration could be significantly reduced with the SHRF, leading to a higher speech intelligibility.

Conclusion

This article addressed the topic of artificial spatialization perception by aided HI subjects. Specifically, a novel feature designed to restore binaural hearing within FM systems was evaluated in terms of speech intelligibility and speaker localization. Several conclusions could be drawn from this study:

- The SHRF did improve the speech understanding of the tested HI subjects presenting a moderate or severe hearing loss, when the FM-transmitted voice was played back alone (FM-only configuration). This means that the conventional full binaural summation operated by current FM systems can be transformed to incorporate the binaural cues required to localize the speaker, without any loss of intelligibility. The advantage obtained with the SHRF was lost with certain subjects when the spatialized speech signal was mixed with the acoustic signals picked up by the HA microphones (FM+M configuration);
- The spatial hearing provided by the SHRF decreased the localization performance of all tested NH listeners, but it is uncertain whether this was due to the use of hearing aids by inexperienced subjects or to the spatial processing itself. No general conclusion could be drawn for the HI subjects, but a majority of them improved or preserved their localization abilities with the spatialization processing. This suggests that HI subjects are less sensitive than NH listeners when hearing with approximated generic HRTFs;
- The human auditory system is able to adapt to hearing impairment to maintain satisfying localization performance up to a certain degree of hearing loss, above which the localization abilities dramatically fall,
- Intelligibility experiments involving severe-to-profound HI listeners should include audiovisual stimuli to allow for lip reading. In this study, such an approach could have highlighted whether the SHRF provides an advantage on the time required

for the speaker identification, and thus on the access to lip reading.

References

Abel, S. M., Gigure, C., Consoli, A., & Papsin, B. C. (2000). The effect of aging on horizontal plane sound localization. *The Journal of Acoustical Society of America*, *108*(2), 743–752. doi: 10.1121/1.429607.

Algazi, V. R., Duda, R. O., Thompson, D. M., & Avendano, C. (2001). The CIPIC HRTF database. *Proceedings of the workshop on the applications of signal processing to audio and acoustics (WASPAA), IEEE* (pp. 99–102). New Paltz, NY: Mohonk Mountain House. doi: 10.1109/ASPAA.2001.969552.

American National Standards Institute. (2003). *ANSI S3.22: Specification of hearing aid characteristics*. Melville, NY: Acoustical Society of America.

Appleton, J., & König, G. (2014). Improvement in speech intelligibility and subjective benefit with binaural beamformer technology. *Hearing Review*, *21*, 40–42.

Arbogast, T. L., Mason, C. R., & Kidd, G., Jr. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *The Journal of Acoustical Society of America*, *117*(4), 2169–2180. doi: 10.1121/1.1861598.

Begault, D., Wenzel, E., Lee, A., & Anderson, M. (2001). Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *108th Audio Engineering Society Convention, 5134*. Paris, France: Audio Engineering Society.

Begault, D. R., & Erbe, T. (1994). Multichannel spatial auditory display for speech communications. *Journal of the Audio Engineering Society*, 42(10), 819–826.

Beskow, J., Dahlquist, M., Granström, B., Lundeberg, M., Spens, K. E., & Öman, T. (1997). The teleface project: Multi-modal speech-communication for the hearing impaired. *Proceedings of Euroseepch '97*. Citeseer. Retrieved from http://www.speech.kth.se/˜beskow/papers/TelefaceES97.pdf.

Best, V., Kalluri, S., McLachlan, S., Valentine, S., Edwards, B., & Carlile, S. (2010). A comparison of CIC and BTE hearing aids for three-dimensional localization of speech. *International Journal of Audiology*, 49(10), 723–732. doi: 10.3109/14992027.2010.484827.

Best, V., Roverud, E., Mason, C. R., & Kidd, G., Jr. (2017). Examination of a hybrid beamformer that preserves auditory spatial cues. *The Journal of Acoustical Society of America*, 142(4), EL369–EL374. doi: 10.1121/1.5007279.

Blauert, J. (1997). Spatial hearing with one sound source. In *Spatial hearing: the psychophysics of human sound localization* (pp. 36–200). Cambridge, MA: MIT press.

Bohnert, A., Nyffeler, M., & Keilmann, A. (2010). Advantages of a non-linear frequency compression algorithm in noise. *European Archives of Otorhinolaryngology*, 267(7), 1045–1053. doi: 10.1007/s00405-009-1170-x.

Boyd, A. W., Whitmer, W. M., Soraghan, J. J., & Akeroyd, M. A. (2012). Auditory externalization in hearing-impaired listeners: The effect of pinna cues and number of talkers. *The Journal of Acoustical Society of America*, 131(3), EL268–EL274. doi: 10.1121/1.3687015.

Bronkhorst, A. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica United With Acustica*, 86(1), 117–128.

Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *The Journal of Acoustical Society of America*, 83(4), 1508–1516. doi: 10.1121/1.395906.

Bronkhorst, A. W., & Plomp, R. (1989). Binaural speech intelligibility in noise for hearing-impaired listeners. *The Journal of Acoustical Society of America*, 86(4), 1374–1383. doi: 10.1121/1.398697.

Brungart, D. S., Cohen, J. I., Zion, D., & Romigh, D. (2017). The localization of non-individualized virtual sounds by hearing impaired listeners. *The Journal of Acoustical Society of America*, 141(4), 2870–2881. doi: 10.1121/1.4979462.

Byrne, D., & Noble, W. (1998). Optimizing sound localization with hearing aids. *Trends in Amplification*, 3(2), 51–73. doi: 10.1177/108471389800300202.

Carhart, R. (1965). Monaural and binaural discrimination against competing sentences. *The Journal of Acoustical Society of America*, 37(6), 1205–1205. doi: 10.1121/1.1939552.

Catic, J., Santurette, S., & Dau, T. (2015). The role of reverberation-related binaural cues in the externalization of speech. *The Journal of Acoustical Society of America*, 138(2), 1154–1167. doi: 10.1121/1.4928132.

Catic, J., Santurette, S., Dau, T., & Buchholz, J. (2012). Effects of interaural level differences on the externalization of sound. *The Journal of Acoustical Society of America*, 131(4), 3520–3520. doi: 10.1121/1.4709317.

Cheng, C. I., & Wakefield, G. H. (1999). Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space. *Proceedings of the Audio Engineering Society Convention 107*. New York, NY: Audio Engineering Society.

Collège National d'Audioprothèse. (2006). *Hearing In Noise Test (HINT) - French database* Retrieved from http://www.college-nat-audio.fr/fichiers/img94a.pdf.

Courtois, G., Marmaroli, P., Lissek, H., Oesch, Y., & Balande, W. (2015a). Binaural hearing aids with wireless microphone systems including speaker localization and spatialization. *Proceedings of the Audio Engineering Society Convention 138*. New York, NY: Audio Engineering Society.

Courtois, G., Marmaroli, P., Lissek, H., Oesch, Y., & Balande, W. (2015b). Development and assessment of a localization algorithm implemented in binaural hearing aids. *23rd European Signal Processing Conference (EUSIPCO)*. Nice, France: European Association for Signal Processing (EURASIP).

Courtois, G., Marmaroli, P., Rohr, L., Lissek, H., Oesch, Y., & Balande, W. (2016). Dynamic range limiting of HRTFs: Principle and objective evaluation. *Journal of the Audio Engineering Society*, 64(10), 731–739.

Crandell, C. J., & Smaldino, J. J. (1999). Improving classroom acoustics: Utilizing hearing-assistive technology and communication strategies in the educational setting. *Volta Review*, 101(5), 47–62.

Cranford, J. L., Andres, M. A., Piatz, K. K., & Reissig, K. L. (1993). Influences of age and hearing loss on the precedence effect in sound localization. *Journal of Speech, Language, and Hearing Research*, 36(2), 437–441. doi: 10.1044/jshr.3602.437.

Culling, J. F., & Mansell, E. R. (2013). Speech intelligibility among modulated and spatially distributed noise sources. *The Journal of Acoustical Society of America*, 133(4), 2254–2261. doi: 10.1121/1.4794384.

Dillon, H. (2012). Binaural and bilateral considerations in hearing aid fitting. *Hearing Aids*, 15, 430–468.

Dirks, D. D., & Wilson, R. H. (1969). The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech, Language, and Hearing Research*, 12(1), 5–38. doi: 10.1044/jshr.1201.05.

Divenyi, P. L., Stark, P. B., & Haupt, K. M. (2005). Decline of speech understanding and auditory thresholds in the elderly. *The Journal of Acoustical Society of America*, 118(2), 1089–1100. doi: 10.1121/1.1953207.

Dobreva, M. S., O'Neill, W., & Paige, G. D. (2011). Influence of aging on human sound localization. *Journal of Neurophysiology*, 105(5), 2471–2486. doi: 10.1152/jn.00951.2010.

Drennan, W. R., Gatehouse, S., Howell, P., Van Tasell, D., & Lund, S. (2005). Localization and speech-identification ability of hearing-impaired listeners using phase-preserving amplification. *Ear and Hear*, 26(5), 461–472. doi: 10.1097/01.aud.0000179690.30137.21.

Drullman, R., & Bronkhorst, A. W. (2000). Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation. *The

Journal of Acoustical Society of America, *107*(4), 2224–2235. doi: 10.1121/1.428503.

Duquesnoy, A. J., & Plomp, R. (1980). Effect of reverberation and noise on the intelligibility of sentences in cases of presbyacusis. *The Journal of Acoustical Society of America*, *68*(2), 537–544. doi: 10.1121/1.384767.

Duquesnoy, A. J., & Plomp, R. (1983). The effect of a hearing aid on the speech reception threshold of hearing impaired listeners in quiet and in noise. *The Journal of Acoustical Society of America*, *73*(6), 2166–2173. doi: 10.1121/1.389540.

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. doi: 10.3109/00206098109072694.

Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *The Journal of Acoustical Society of America*, *106*(6), 3578–3588. doi: 10.1121/1.428211.

Gallun, F. J., Mason, C. R., & Kidd, G., Jr. (2005). Binaural release from informational masking in a speech identification task. *The Journal of Acoustical Society of America*, *118*(3), 1614–1625. doi: 10.1121/1.1984876.

George, E. J. L., Festen, J. M., & Houtgast, T. (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *The Journal of Acoustical Society of America*, *120*(4), 2295–2311. doi: 10.1121/1.2266530.

Glasberg, B. R., & Moore, B. C. J. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, *50*(5), 331–342.

Hassager, H. G., May, T., Wiinberg, A., & Dau, T. (2017a). Preserving spatial perception in rooms using direct-sound driven dynamic range compression. *The Journal of Acoustical Society of America*, *141*(6), 4556–4566. doi: 10.1121/1.4984040.

Hassager, H. G., Wiinberg, A., & Dau, T. (2017b). Effects of hearing-aid dynamic range compression on spatial perception in a reverberant environment. *The Journal of Acoustical Society of America*, *141*(4), 2556–2568. doi: 10.1121/1.4979783.

Hawkins, D. B. (1984). Comparisons of speech recognition in noise by mildly-to-moderately hearing-impaired children using hearing aids and FM systems. *The Journal of Speech and Hearing Disorders*, *49*(4), 409–418. doi: 10.1044/jshd.4904.409.

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *The Journal of Acoustical Society of America*, *115*(2), 833–843. doi: 10.1121/1.1639908.

Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of Acoustical Society of America*, *130*(1), 334–349. doi: 10.1121/1.3585848.

Keidser, G., O'Brien, A., Hain, J. U., McLelland, M., & Yeend, I. (2009). The effect of frequency-dependent microphone directionality on horizontal localization performance in hearing-aid users. *International Journal of Audiology*, *48*(11), 789–803. doi: 10.3109/14992020903036357.

Keidser, G., Rohrseitz, K., Dillon, H., Hamacher, V., Carter, L., Rass, U., & Convery, E. (2006). The effect of multichannel wide dynamic range compression, noise reduction, and the directional microphone on horizontal localization performance in hearing aid wearers. *International Journal of Audiology*, *45*(10), 563–579. doi: 10.1080/14992020600920804.

Kreisman, B. M., Mazevski, A. G., Schum, D. J., & Sockalingam, R. (2010). Improvements in speech understanding with wireless binaural broadband digital hearing instruments in adults with sensorineural hearing loss. *Trends in Amplification*, *14*(1), 3–11. doi: 10.1177/1084713810364396.

Latzel, M. (2013). Concepts for binaural processing in hearing aids. *Hearing Review*, *20*(4), 34.

Lewis, M. S., Crandell, C. C., Valente, M., & Horn, J. E. (2004). Speech perception in noise: Directional microphones versus frequency modulation (FM) systems. *Journal of the American Academy of Audiology*, *15*(6), 426–439. doi: 10.3766/jaaa.15.6.4.

Macleod, A., & Summerfield, Q. (1990). A procedure for measuring auditory and audiovisual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British Journal of Audiology*, *24*(1), 29–43. doi: 10.3109/03005369009077840.

Majdak, P., Walder, T., & Laback, B. (2013). Effect of long-term training on sound localization performance with spectrally warped and band-limited head-related transfer functions. *The Journal of Acoustical Society of America*, *134*(3), 2148–2159. doi: 10.1121/1.4816543.

McCreery, R. W., Alexander, J., Brennan, M. A., Hoover, B., Kopun, J., & Stelmachowicz, P. G. (2014). The influence of audibility on speech recognition with nonlinear frequency compression for children and adults with hearing loss. *Ear and Hear*, *35*(4), 440–447. doi: 10.1097/AUD.0000000000000027.

Mendonca, C., Campos, G., Dias, P., Vieira, J., Ferreira, J. P., & Santos, J. A. (2012). On the improvement of localization accuracy with non-individualized HRTF-based sounds. *Journal of the Audio Engineering Society*, *60*(10), 821–830.

Minnaar, P. (2010). Enhancing music with virtual sound sources. *The Hearing Journal*, *63*(9), 38–40. doi: 10.1097/01.HJ.0000388539.35029.2c.

Moore, B. C. J. (2007). Spatial hearing and advantages of binaural hearing. In B. C. J. Moore (Ed.), *Cochlear hearing loss: Physiological, psychological and technical issues* (pp. 173–200). Hoboken, NJ: John Wiley & Sons.

Moore, B. C. J., Peters, R. W., & Stone, M. A. (1999). Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips. *The Journal of Acoustical Society of America*, *105*(1), 400–411. doi: 10.1121/1.424571.

Neher, T. (2017). Characterizing the binaural contribution to speech-in-noise reception in elderly hearing-impaired listeners. *The Journal of Acoustical Society of America*, *141*(2), EL159–EL163. doi: 10.1121/1.4976327.

Neher, T., Wagener, K. C., & Latzel, M. (2017). Speech reception with different bilateral directional processing schemes:

Influence of binaural hearing, audiometric asymmetry, and acoustic scenario. *Hearing Research*, *353*, 36–48. doi: 10.1016/j.heares.2017.07.014.

Noble, W., & Byrne, D. (1990). A comparison of different binaural hearing aid systems for sound localization in the horizontal and vertical planes. *British Journal of Audiology*, *24*(5), 335–346. doi: 10.3109/03005369009076574.

Noble, W., Byrne, D., & Ter-Horst, K. (1997). Auditory localization, detection of spatial separateness, and speech hearing in noise by hearing impaired listeners. *The Journal of Acoustical Society of America*, *102*(4), 2343–2352. doi: 10.1121/1.419618.

Nordlund, B. (1964). Directional audiometry. *Acta Oto-Laryngologica*, *57*(1–2), 1–18. doi: 10.3109/00016486409136942.

Ohl, B., Laugesen, S., Buchholz, J., & Dau, T. (2010). Externalization versus internalization of sound in normal-hearing and hearing-impaired listeners. In: *Tagung der Deutschen Arbeitsgemeinschaft fr Akustik (DAGA)*. Deutsche Gesellschaft fur Akustik (DEGA). Retrieved from http://orbit.dtu.dk/en/publications/externalization-versus-internalization-of-sound-in-normalhearing-and-hearingimpaired-listeners(fd8b42cf-0a50-4c0e-9316-d5a18ccc65ca).html.

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. doi: 10.1097/AUD.0000000000000004.

Plomp, R., & Mimpen, A. M. (1979). Speech-reception threshold for sentences as a function of age and noise level. *The Journal of Acoustical Society of America*, *66*(5), 1333–1342. doi: 10.1121/1.383554.

Raake, A., & Katz, B. F. G. (2006). SUS-based method for speech reception threshold measurement in French. *Fifth conference on language resources and evaluation (LREC)*, Genoa, Italy, pp. 2028–2033.

Schafer, E. C., Mathews, L., Mehta, S., Hill, M., Munoz, A., Bishop, R., & Moloney, M. (2013). Personal FM systems for children with autism spectrum disorders (ASD) and/or attention-deficit hyperactivity disorder (ADHD): An initial investigation. *Journal of Communication Disorders*, *46*(1), 30–52. doi: 10.1016/j.jcomdis.2012.09.002.

Smoorenburg, G. F. (1992). Speech reception in quiet and in noisy conditions by individuals with noiseinduced hearing loss in relation to their tone audiogram. *The Journal of Acoustical Society of America*, *91*(1), 421–437. doi: 10.1121/1.402729.

Sockalingam, R., Holmberg, M., Eneroth, K., & Shulte, M. (2009). Binaural hearing aid communication shown to improve sound quality and localization. *The Hearing Journal*, *62*(10), 46–47. doi: 10.1097/01.HJ.0000361850.27208.35.

Thibodeau, L. (2010). Benefits of adaptive FM systems on speech recognition in noise for listeners who use hearing aids. *American Journal of Audiology*, *19*(1), 36–45. doi: 10.1044/1059-0889(2010/09-0014).

Thibodeau, L. (2014). Comparison of speech recognition with adaptive digital and FM remote microphone hearing assistance technology by listeners who use hearing aids. *American Journal of Audiology*, *23*(2), 201–210. doi: 10.1044/2014_AJA-13-0065.

Thiemann, J., Mller, M., Marquardt, D., Doclo, S., & Van de Par, S. (2016). Speech enhancement for multimicrophone binaural hearing aids aiming to preserve the spatial auditory scene. *Journal on Advances in Signal Processing*, *2016*(1), 12. doi: 10.1186/s13634-016-0314-6.

Timmer, B. (2013). Its sync or stream! The differences between wireless hearing aid features. *Hearing Review*, *20*(6), 20–22.

Tremblay, K. L., Piskosz, M., & Souza, P. (2003). Effects of age and age-related hearing loss on the neural representation of speech cues. *Clinical Neurophysiology*, *114*(7), 1332–1343. doi: 10.1016/S1388-2457(03)00114-7.

Udesen, J., Piechowiak, T., & Gran, F. (2015). The effect of vision on psychoacoustic testing with headphone-based virtual sound. *Journal of Audio Engineering Society*, *63*(7/8), 552–561.

Van den Bogaert, T., Carette, E., & Wouters, J. (2009). Sound localization with and without hearing aids. *International Conference on Acoustics (NAG-DAGA)* (pp. 1–4). Berlin, Germany: German Acoustical Society (DEGA).

Van den Bogaert, T., Klasen, T. J., Moonen, M., Van Deun, L., & Wouters, J. (2006). Horizontal localization with bilateral hearing aids: Without is better than with. *The Journal of Acoustical Society of America*, *119*(1), 515–526. doi: 10.1121/1.2139653.

Wenzel, E. M., Arruda, M., Kistler, D. J., & Wightman, F. L. (1993). Localization using nonindividualized head-related transfer functions. *The Journal of Acoustical Society of America*, *94*(1), 111–123. doi: 10.1121/1.407089.

Whitmer, W. M., Seeber, B. U., & Akeroyd, M. A. (2014). The perception of apparent auditory source width in hearing-impaired adults. *The Journal of Acoustical Society of America*, *135*(6), 3548–3559. doi: 10.1121/1.4875575.

World Health Organization. (2016). Prevention of blindness and deafness. Retrieved from http://www.who.int/pbd/deafness/hearing_impairment_grades/en/

Wiggins, I. M., & Seeber, B. U. (2012). Effects of dynamic-range compression on the spatial attributes of sounds in normal-hearing listeners. *Ear and Hearing*, *33*(3), 399–410. doi: 10.1097/AUD.0b013e31823d78fd.

Wightman, F. L., & Kistler, D. J. (1989a). Headphone simulation of free-field listening. I: Stimulus synthesis. *The Journal of Acoustical Society of America*, *85*(2), 858–867. doi: 10.1121/1.397557.

Wightman, F. L., & Kistler, D. J. (1989b). Headphone simulation of free-field listening. II: Psychophysical validation. *The Journal of Acoustical Society of America*, *85*(2), 868–878. doi: 10.1121/1.397558.

Yousefian, N., Loizou, P. C., & Hansen, J. H. L. (2014). A coherence-based noise reduction algorithm for binaural hearing aids. *Speech Communication*, *58*, 101–110. doi: 10.1016/j.specom.2013.11.003.