Inform Health Soc Care, 2014; 39(3–4): 188–209

ISSN: 1753-8157 print / 1753-8165 online
DOI: 10.3109/17538157.2014.931856

informa
healthcare

Audio reproduction for personal ambient home assistance: concepts and evaluations for normal-hearing and hearing-impaired persons

Rainer Huber,[1,2] Markus Meis,[2,3] Karin Klink,[2,4]
Christian Bartsch,[2,5] and Joerg Bitzer[2,5,6]

[1]Center of Competence HörTech gGmbH, Oldenburg, Germany,
[2]Cluster of Excellence Hearing4all, Oldenburg, Germany,
[3]Hörzentrum Oldenburg GmbH, Oldenburg, Germany,
[4]Medizinische Physik, Universität Oldenburg, Oldenburg, Germany,
[5]HSA, Fraunhofer IDMT, Oldenburg, Germany, and
[6]IHA, Jade University of Applied Science, Oldenburg, Germany

Within the Lower Saxony Research Network Design of Environments for Ageing (GAL), a personal activity and household assistant (PAHA), an ambient reminder system, has been developed. One of its central output modality to interact with the user is sound. The study presented here evaluated three different system technologies for sound reproduction using up to five loudspeakers, including the "phantom source" concept. Moreover, a technology for hearing loss compensation for the mostly older users of the PAHA was implemented and evaluated. Evaluation experiments with 21 normal hearing and hearing impaired test subjects were carried out. The results show that after direct comparison of the sound presentation concepts, the presentation by the single TV speaker was most preferred, whereas the phantom source concept got the highest acceptance ratings as far as the general concept is concerned. The localization accuracy of the phantom source concept was good as long as the exact listening position was known to the algorithm and speech stimuli were used. Most subjects preferred the original signals over the pre-processed, dynamic-compressed signals, although processed speech was often described as being clearer.

Keywords Ageing society, acoustic interfaces, ambient-assisted living, hearing impairment, multi-channel audio reproduction

INTRODUCTION

In view of current and expected future demographic changes showing a significant growth of the older population, the exploitation of modern information and communication technologies in the design of supportive home environments ("ambient assisted living") for ageing appears promising to help meeting the associated challenges, e.g. in the field of social care. However, apart from "just" developing solutions for various technical challenges, the successful application of any "intelligent" IT-based assistive

Correspondence: Rainer Huber, Center of Competence HörTech gGmbH, Marie-Curie-Str. 2, D-26129 Oldenburg, Germany. E-mail: Rainer.Huber@HoerTech.de

systems also requires to consequently take into account user needs, user acceptance, the system's integration into medical and nursing care structures and economical aspects. In a number of publications in this journal, the Lower Saxony Research Network Design of Environments for Ageing (GAL) has been introduced, which strives to identify, (further) develop and evaluate such (new) information and communication technologies for ageing while addressing all of the mentioned issues (see, e.g. the opening article in this special issue).

One of the assistive scenarios developed and evaluated in the GAL project is a personal activity and household assistant (PAHA), an ambient reminder system, which uses various input and output modalities (1,2). Next to visual and tactile outputs, sound, including speech, represents a central output modality.

The focus of the study presented in this article will be on the evaluation of system technologies for audio reproduction using loudspeakers. In the GAL context, acoustical signal presentation can be used for conveying various kinds of information to the user. Examples are alarm sounds for warning the user that electrical kitchen devices have not been switched off, that the refrigerator door has not been closed, or signaling sounds that remind the user of appointments or tasks. Moreover, acoustical output is used in speech dialogues between user and electronic calendar for programming the calendar. Apart from that, the installed loudspeakers can be used in other, common multimedia applications like home cinema. They might even already be installed in the users' apartment for such purposes anyway; in that case, the existing equipment could be used to save costs. The simplest system technology for acoustic signal presentation is playing the sound with single loudspeaker (mono presentation). If two loudspeakers are available, a stereo sound pattern can be generated, which could be optimized for specific spatial listening positions (3). Another technology employing two loudspeakers allows for manipulating the perceived direction of a (virtual) sound source (4,5). This so-called *phantom source* concept is restricted, though, in such a way that the perceived position of the virtual sound source can only be located *between* the two loudspeakers. Since the phantom source concept is not restricted to a certain number of loudspeakers, this restriction can be overcome by simply adding more loudspeakers that surround the area of desired virtual sound source locations. Moreover, using several loudspeakers altogether allows for the possibility to generate simple spatial surround sounds impressions. More sophisticated methods using arrays of many loudspeakers are higher-order ambisonics (HOA) (6,7) and wave field synthesis (WFS) (8). In WFS, a large number of closely spaced loudspeakers are used. Following the Huygens principle, the superposition of many widely overlapping "elementary waves" builds the resulting sound wave. Both of these concepts aim at approximating the actual physical sound pressure field within a certain listening area, which can be significantly larger than typical sizes of the "sweet spots" of conventional stereo or 5.1 surround systems commonly used in multimedia home entertainment systems. However, such concepts, wave field synthesis in particular, appear too complex and costly for simple home applications.

Consequently, three different sound presentation concepts of increasing complexity using up to five loudspeakers have been selected for implementation and evaluation in the context of the PAHA in the GAL project, which will be presented in detail in the following sections.

Another aspect of sound presentation considered in the present study was the compensation of possible hearing losses by appropriate pre-processing of audio signals. About half of the people aged 65 and above have hearing impairments (9,10), yet studies estimate that only about 30% of those people are actually aided with hearing aids (11). Since the target group of the GAL system technology are mainly older people (55+), it can be assumed that a significant portion of the users has hearing impairments, but does not (yet) wear hearing aids and could thus benefit from hearing-aid-like signal pre-processing. Because interaction with the electronic calendar of the PAHA system is based on speech dialogues, deteriorated speech intelligibility due to hearing impairments can be a critical issue. Signal processing strategies for the compensation of hearing losses is thus deemed an important component of the sound presentation concept of the PAHA system. Consequently, a corresponding hearing-aid-like signal pre-processing (multi-band dynamic compression) was implemented in the sound playback module of the PAHA system and tested in the present study as well.

Hearing impairments can also have detrimental effects on the ability to localize sounds. This also applies to aided hearing impaired people wearing only one hearing aid (i.e. unilaterally) or non-coupled hearing aids on both sides (i.e. bilaterally, in contrast to "binaurally"), which can distort the original binaural localization cues (i.e. interaural time and level differences). To the knowledge of the authors, the localization accuracy of the phantom source concept had not been tested before with hearing impaired listeners and was thus posed as a research question on its own in the present study.

CONCEPTS AND ALGORITHMS FOR SOUND REPRODUCTION

Loudspeaker presentation concepts

In order to test three signal presentation methods as the audio output of the personal activity and household assistant (PAHA), five loudspeakers are distributed in the room, four of them in the corners and one at the TV set (Figure 4).

We tested the following concepts:

(1) **Single speaker output**: Only the loudspeaker near the TV set is used as the output device. The advantage for this method is the defined position of the source; a user could adapt and learn to listen in the well-defined direction. For a single channel method, the TV set is the natural choice, since the TV is one key component for the overall concept of the PAHA.

(2) **All unisono**: All loudspeakers playing the same audio signal for the impression of an omnipresent system. This can be implemented by using individual gains and delays for each loudspeaker to compensate for equal loudness and equal time of arrival at the position of the listener. Without this procedure, only the nearest loudspeaker would be perceived

as the dominating source. The time delay for each loudspeaker is computed by

$$T_i = \frac{d_i}{c}$$

where c denotes the speed of sound and d_i the distance between the loudspeaker and the position of the user. The maximum of all T_i is used as the reference and all other signals are delayed to get the same time of arrival at the listener's position.

$$\Delta T_i = \max(T_i) - T_i$$

The gains are set according to (12) to incorporate room acoustics:

$$a_i = \frac{e^{-r \cdot \Delta T_i}}{1 + c \cdot \Delta T_i / d_i}$$

where

$$r = \frac{\ln(1000)}{T_{60}}$$

denotes the absorption coefficient and T_{60} is the reverberation time measured as 0.365 s in the room.

(3) **Phantom sources**: The idea of our third method is to combine the content of the message with the appropriate direction, for example a door bell could be assigned in the direction of the door and a reminder to switch off the stove in the direction of the kitchen.

The concept of phantom sources is well known for stereo music processing (13). A virtual source can be generated if two spatially separated loudspeakers are playing the same signal. If the listener is symmetrically in front of the two loudspeakers, a virtual source in the center of the two loudspeakers will be perceived. In our case the distance and the involved positions are not symmetrical. Therefore, we applied the methods of Pullki (4) and the optimization of Van Leest (5) in order to get the necessary gains and delays for optimal positions of the virtual source.

Hearing loss compensation

Independent of the three loudspeaker presentation concepts described above, test stimuli for two of the three subject groups (i.e. the ones without hearing aids) were pre-processed optionally by a multi-band dynamic compression algorithm running on HörTech's Master Hearing Aid (MHA (14)). (The MHA has been integrated into the general GAL system platform.) The aim of this frequency- and input-level-dependent amplification typically used in modern digital hearing aids is to compensate for a possible hearing loss of the user. In the present evaluation, the dynamic compression algorithm gains were set to match the average hearing loss per subject group following the NAL NL1 fitting rule (15). Further, hearing-loss-independent parameters were set as follows:

- Number of independent frequency channels: 9

- Channel center frequencies: 250, 500, 1000, 1500, 2000, 3000, 4000, 6000, 8000 Hz
- Attack time: 10 ms (the same in all frequency channels)
- Release time: 15 ms (the same in all frequency channels)

EVALUATION

Aim and design of the study

The aim of the evaluation study was to determine

(1) the acceptance and preference of the three sound presentation concepts
(2) the localization accuracy of the phantom source algorithm
(3) whether a hearing loss compensation by signal pre-processing is preferred over no processing

For this purpose, a number of young-old to middle-old test subjects with normal hearing and mild to moderate hearing losses (with and without hearing aids) were to perform localization tests in a home-like test environment and to rate the acceptance and preference of the sound presentation concepts by means of questionnaires after the concepts have been demonstrated. In the following, test subjects, setup and methods will be described in detail.

Methods

Test subjects

In total, 21 test subjects participated in the evaluation study. Since this study represents the first investigation of the described research question, there was no prior knowledge or expectation regarding effect size, mean values or standard deviations that could have been used as a basis to calculate a required sample size. Hence, the chosen number of test subjects represents a typical number of subjects employed in earlier studies in related fields of research of the authors' institutions, where this sample size has been found to be reasonable in most cases. The subjects have been recruited from the test subject database of the project partner Hörzentrum Oldenburg. One criterion for selecting the subjects was their participation in earlier studies of the GAL project or others projects. This criterion was met for most of the subjects. In those earlier studies, the subjects' cognitive and motor performances appeared to be age-appropriate or better.

The other main criterion was the kind and degree of hearing loss and the fact whether they wear hearing aids nor not. The aim was to compose a well selected, representative sample of potential users of the tested system. According to these criterions, three groups of 7 subjects each were recruited:

- Group 1: Normal hearing according to WHO (16) (NH)
- Group 2: Mildly to moderately hearing impaired, unaided, i.e. without hearing aid (HI-HA)
- Group 3: Mildly to moderately hearing impaired, aided (HI+HA)

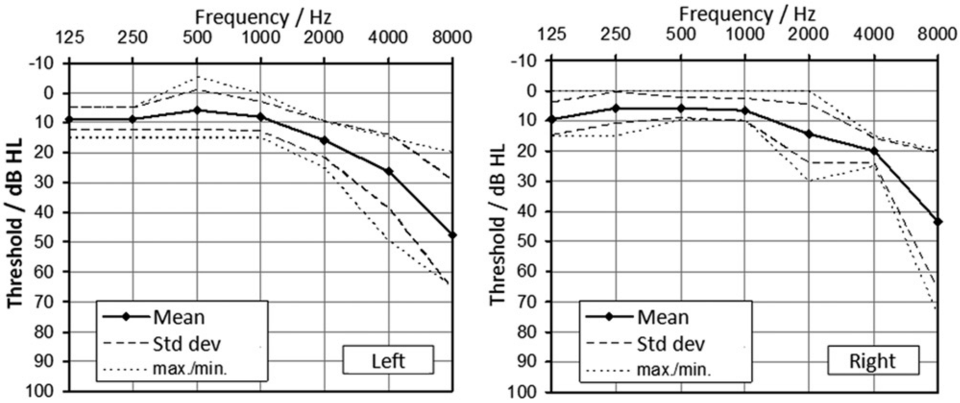

Figure 1. Audiograms of the normal-hearing subject group (NH). Left: left ear; right: right ear.

All subjects participated in the localization tests as well in the tests to determine the preference and acceptance of the sound presentation concepts. Groups 1 (NH) and 2 (HI-HA), i.e. all unaided subjects, also participated in the test to determine the preference/non-preference for pre-processed audio signals (hearing loss compensation).

These three groups will be characterized in detail in the following.

Group 1: Normal hearing (NH). The group of the (age-according) normal hearing subjects consisted of five female and two male aged between 59 and 72 years (mean = 66.7, standard deviation = ±5.9). The group average of the pure tone average hearing threshold (PTA) of the better ear in the frequency range of 0.5 kHz – 4 kHz ("Better Ear Hearing Loss" – BEHL) was 10.7 dB HL ± 3.7 dB HL. According to the WHO criterion (BEHL < 25 dB HL), subjects with such hearing loss are considered normal hearing (16). The audiograms of group 1 are summarized in Figure 1.

Group 2: Hearing impaired, unaided (HI-HA). The group of subjects with mild-to-moderate hearing losses (according to WHO (16)) without hearing aids contained three male and four female, aged between 61 and 74 years (68.9 ± 4.5). The group average BEHL was 31.1 ± 1.7 dB HL. All hearing losses were typical, sensorineural hearing losses, basically symmetrical between left and right ear. The distribution of hearing losses in terms of pure-tone audiograms was less homogeneous in this group than in the other groups; the standard deviation between individual audiograms (averaged across all frequencies and sides) was 15.2 dB HL, which is almost twice as large as those of the other two subject groups (7.9 dB HL) (Figure 2).

Group 3: Hearing impaired, aided (HI+HA). Five women and two men with moderate hearing losses (16), wearing hearing aids, constituted the third subject group HI+HA. The age span was 66 to 74 years (69.4 ± 2.9). The group's mean BEHL was 40.5 ± 6.2 dB HL. Again, hearing losses were typical, sensorineural and left/right symmetric (Figure 3).

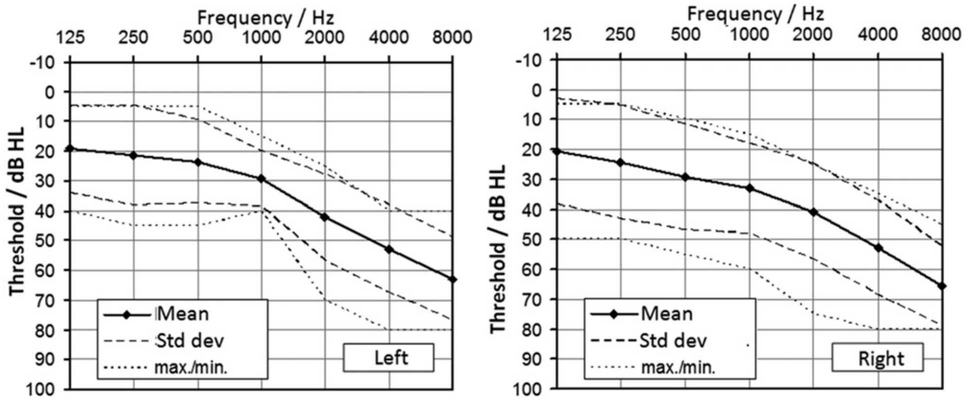

Figure 2. Audiograms of the unaided hearing impaired subject group (HI-HA) with mild-to-moderate hearing losses. Left: left ear; right: right ear.

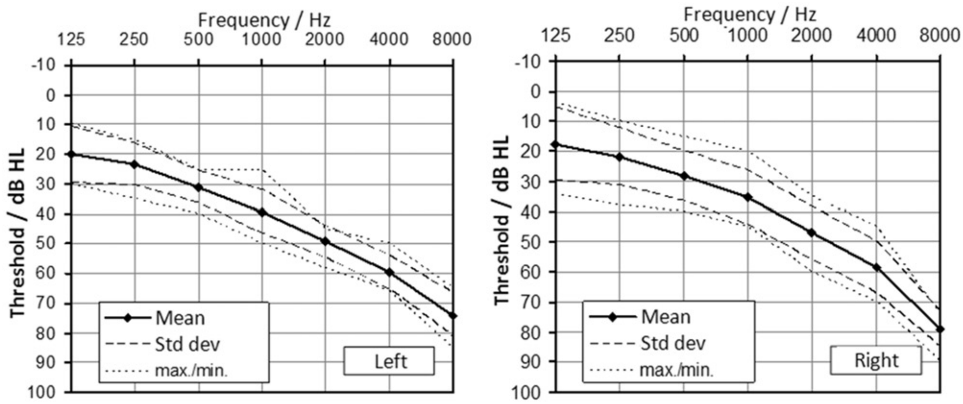

Figure 3. Audiograms of the aided hearing impaired subject group (HI+HA) with moderate hearing losses. Left: left ear; right: right ear.

Measurement set-up: Spatial configurations

All experiments have been carried out in the living room of the IDEAAL apartment, i.e. a home-lab for studies in the context of ambient assisted living (AAL) at the OFFIS institute in Oldenburg. The room was equipped with an array of five loudspeakers, positioned at head-hight of the (sitting) test subjects as indicated by crosses in Figure 4 (See this figure also for size and geometry of the test room).

The experiments were carried out for three different listener positions (1a, 1b and 2; see corresponding labels in Figure 4). In positions 1a and 1b, the test subject sat in an armchair near the centre of the room. In position 1a, listeners bended forward, whereas in position 2a, they leaned against the back of the armchair. (The height of the back rest was lower than the shoulders of the subjects, so it did not interfere with the sound field at the subject's ear positions.) In position 2, listeners sat on a sofa near one of the walls. When testing the phantom source sound presentation concept, the listener's position was assumed to be known and provided to the algorithm as a constant parameter. (In future implementations of the complete GAL system, the user's

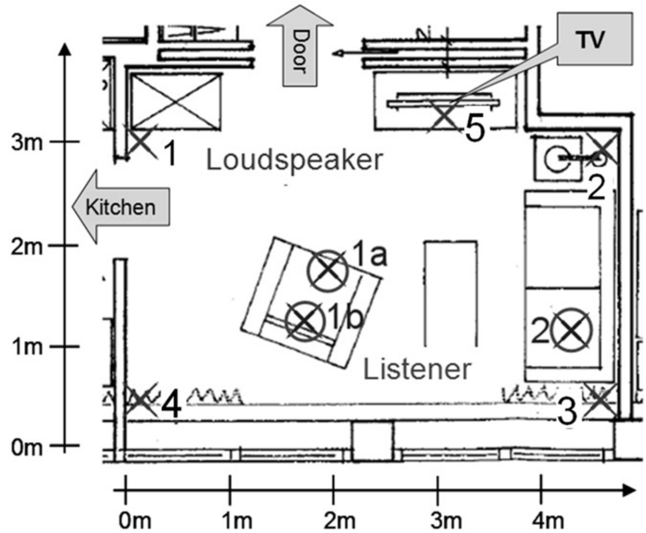

Figure 4. Sketch of the IDEAAL living room with markers for loudspeaker positions (crosses near the walls) and test subject positions (''Listener'', encircled crosses).

actual position could be determined by localization techniques developed in other work packages of the GAL project and provided to the sound presentation module.) In the phantom source algorithm, the coordinates of position 1a were used for both actual positions 1a and 1b, leading to a sub-optimal setting of the algorithm in case of the actual listener position 1b. This mismatch of about 0.5 m was set intentionally to investigate the susceptibility of the phantom source algorithm to inaccuracies of (future) automatic user localization techniques. At position 2, distances between listener and the five different loudspeakers differed significantly (0.85–4.30 m). This configuration represents a more critical and demanding test for the multi-loudspeaker presentation concepts "phantom source" and "all loudspeakers".

Apparatus

The playback of audio stimuli as well as the complete measurement was controlled by a self-made MATLAB software module, running on a Windows laptop. The software was controlled (via graphical user interface) by the experimenter, who was seated in the IDEAAL living room together with the test subject. (The presence of the experimenter in the measurement room was mainly required to collect the subject's responses.) Audio stimuli were put out digitally by an external, 8-channel sound card (RME Digiface) and converted to analog signals by an external D/A converter (Behringer ULTRAGAIN PRO-8 Digital). The analog audio signals were amplified by an OMNITRONIC MPS-1250 multi-channel amplifier before emitted by five Bosch LB1-CW06-L loudspeakers. Soundcard, D/A converter and amplifier were located outside the measurement room. Acoustic level calibration was done by means of the software module and a level meter (NTI Acoustilyzer AL1).

Test stimuli

In the localization experiment, three different acoustic stimuli were used: [1] speech, [2] click train (rate: 100/s), [3] gong. All signals had similar durations (about 4 s) and (comfortable) loudnesses. Three acoustic stimuli were also used in the preference experiment, corresponding to the simulated event (see section “Preference determination”): [1] speech, [2] door bell, [3] synthetic alarm sound. The sample rate of all signals was 44.1 kHz.

Measurement procedures

Prior to the experiments, the background of the study (i.e. the overall aim of the GAL project) and the purpose of the current study and of each experiment in particular were explained to the subjects in oral and in written form.

Localization experiment. In each of the three listener positions 1a, 1b and 2 (cf. Figure 4), the three acoustic stimuli speech, click train and gong were played back to the test subjects from three different phantom source directions ([1] kitchen, [2] door, and [3] TV) using the phantom source algorithm. This gave a total of $3 \times 3 = 9$ sound presentations per listening position. The order of listening positions, source direction and type of test stimulus was randomized across subjects. At the request of the subjects, the presentation of sound stimuli was repeated until the subject got to a clear conclusion about the source direction. Subjects indicated the perceived directions of the incident sounds by pointing into the corresponding direction with a laser pointer; a common, validated method in acoustic localization experiments (17,18). The horizontal position of the laser spot on the walls was measured and logged by the experimenter. For this purpose, measuring tapes had been attached to the surrounding walls. The directions indicated by the subjects with a laser pointer were translated from horizontal positions of the laser spot on the walls into polar coordinates. Localization errors were computed in terms of differences between angles of directions intended by the phantom source algorithm and angles of perceived directions indicated by the subjects. In the following, positive errors mean that the perceived direction was to the right of the “actual”, i.e. intended direction of the phantom source, negative errors indicate deviations to the left-hand side.

Preference determination. In this part of the experiment, subjects’ preferences for one of the three sound presentation concepts

(1) all loudspeakers *unisono*
(2) single (TV) loudspeaker
(3) phantom source

depending on the situation or event were to be determined. To this end, three fictitious events were described to the subjects and simulated acoustically by presenting corresponding test stimuli (cf. section “Test stimuli”):

(a) Date reminder by the electronic calendar of the PAHA (acoustic stimulus: speech)
(b) A visitor ringing the door bell (acoustic stimulus: door bell)
(c) Alarm of a kitchen device (e.g. electric oven or open fridge; acoustic stimulus: synthetic, pulsed alarm sound)

For each of these simulated events, the three described sound presentation concepts were demonstrated to the subjects acoustically. After the last demonstration, the subject was asked to indicate the concept he or she would prefer. In total, nine preference assessments were made by each subject. (Three listening positions times three events.)

Acceptance assessment of the concepts. After the preference determination experiment, the acceptance of the three sound presentation concepts [1] identical sound play-back by all loudspeakers distributed in the room or the apartment, respectively, [2] central play-back by the TV loudspeaker in general, [3] play-back from a (perceived) situation-dependent direction ("acoustical guide post") was assessed by means of a questionnaire. The concepts have been described and explained to the subjects once more by the experimenter, before the subjects were asked to rate their acceptance for the respective concepts on a five-point scale, ranging from "very little" to "very high". The acceptance was asked in different contexts: With regard to "situations with reference to locations at home" (e.g. door bell or alarm of a kitchen device) and "situations *without* reference to locations at home" (e.g. date reminder). Subjects were asked to state reasons for their rating briefly.

Preference for dynamic-compressed versus original signals. Subjects without hearing aids (i.e. members of the groups NH and HI-HA) were asked whether they prefer hearing loss compensation by signal pre-processing over no processing. The preprocessing consisted of a multiband dynamic compression used in modern hearing aids. The dynamic compression parameters were set to compensate for the average hearing loss of the respective subject group using a standard prescription rule for hearing aids. The algorithm was hosted on HörTech's Master Hearing Aid (14). In order to assess and compare the processing/no processing conditions, the same three acoustic stimuli (A, B, C) as used in the preference determination experiment were played back once more, with and without dynamic compression. Listening position 1a was used. In each paired comparison, the presentation order of the unprocessed and processed audio signal was randomized. After each pair of presentation, subjects were asked to indicate which of the two variants they would prefer in the everyday usage of the system.

Statistical analyses methods

Prior to the actual analyses, the localization experiment data were tested positively for normal distribution using the Kolmogorov–Smirnov test, which justified the application of parametric analyses in a model. Following the General Linear Model (GLM), an analysis of variance was computed with "within" factors listening position (1a, 1b, 2) × sound source direction (door, TV, kitchen) × test stimuli (speech, click-train, gong) and the "between" factor subject group (NH, HI-HA, HI+HA). A Bonferoni correction was applied to the post-hoc tests concerning repeated comparisons. Repeated-measure computations of main effects were corrected by the Greenhouse–Geisser correction method (19,20). The advantage of this approach is the computation with a single variance-analytic model, without having to compute individual comparisons separately. Level of significance was defined at $p < 0.05$. Statistical analyses were computed with the statistics software SPSS ver. 11.5 (Chicago, IL). In addition to preference and scale ratings, open questions were also asked to justify the ratings; some basic answers will be stated.

RESULTS

Localization experiment

In order to test the accuracy and reliability of the localization of phantom sources, three independent variables were varied: [1] The listening position (1a, 1b, 2), [2] the direction of the phantom source (kitchen, door, TV), and [3] the test stimulus (speech, click train and gong). Statistical analyses of the direction errors using a General Linear Model showed that the localization error magnitude (as a measure for the inverse of the accuracy) depended on listening position, direction of the phantom sound source and type of acoustic stimulus significantly. In contrast, the degree of hearing loss had no significant influence on the localization performance. The results will be presented in detail in the following.

Influence of the listening position

The listening position was found to have a significant influence on the localization accuracy. Subjects located at listening position 1a and 2 could localize the phantom source relatively well (cf. Table 1 and Figure 5); there was no systematic error (ε) to either side (mean localization error over all subjects: 5° and 1°, respectively), the standard deviation of the error (which is equal to the standard deviation of the measured directions) amounted to ±32° and ±27°, respectively. Errors were significantly larger in listening position 1b, with a systematic offset to the left hand side ($\varepsilon = -17° \pm 37°$). Moreover, the observed mean absolute localization errors at positions 1a and 2 were significantly smaller than chance level in all except one of 54 cases, whereas no significant smaller errors than chance level were reached in 11 of 27 cases at listening position 1b.

Influence of the phantom sound source direction

The localization accuracy depended significantly on the direction of the phantom sound source. Sounds intended to originate from the direction of the kitchen were systematically (i.e. on average) localized by 11° to the right of the "actual" direction. The localization of sounds "actually" coming from the direction of the door and the TV were biased into to the opposite (i.e. left) direction by −8° and −14°, respectively. Standard deviations amounted to ±20°

Table 1. Localization errors (in degree) for different listening positions.

Listening position	1a	1b	2
Mean error	5°	−17°	1°
Standard deviation	±32°	±37°	±27°

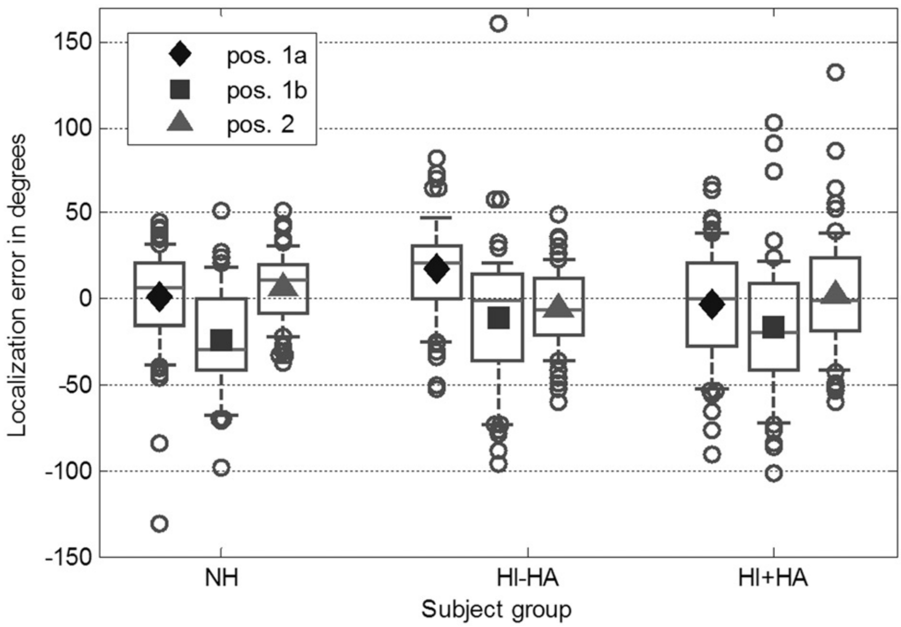

Figure 5. Box plot of localization errors of the phantom source presentation for the three subject groups NH, HI-HA, and HI+HA and three listening positions 1a (diamonds) 1b (squares) and 2 (triangles). Square, diamond and triangle symbols also indicate mean values. Upper and lower boundaries of the boxes represent upper and lower quartiles, respectively, upper and lower whiskers at 91th and 9th percentiles, respectively. Horizontal bars within boxes are median values. Circles indicate outliers (i.e. values larger than the 91th/lower than the 9th percentile, respectively).

Table 2. Localization errors (in degree) for different phantom source positions.

Source position	Door	TV	Kitchen
Mean error	−8°	−14°	11°
Standard deviation	±36°	±38°	±20°

(kitchen), ±36° (door) and ±38° (TV), respectively. (Due to these large stand deviations, mean localized positions did not differ significantly from "actual" phantom source positions.) The localization errors are summarized in Table 2 and Figure 6.

Influence of the acoustic stimulus

The kind of the acoustic stimulus (speech, click train, gong) used in the localization test had a significant effect on the localization accuracy. Speech was connected with the smallest localization errors ($\varepsilon = 3° \pm 28°$), whereas click train and gong lead to somewhat higher deviations of perceived from "actual" directions (click train: $\varepsilon = -8° \pm 35°$; gong: $\varepsilon = -6° \pm 38°$). The localization accuracies with speech and click trains differed significantly, whereas the differences between gong and either of the other signals did not reach level of significance (cf. Figure 7).

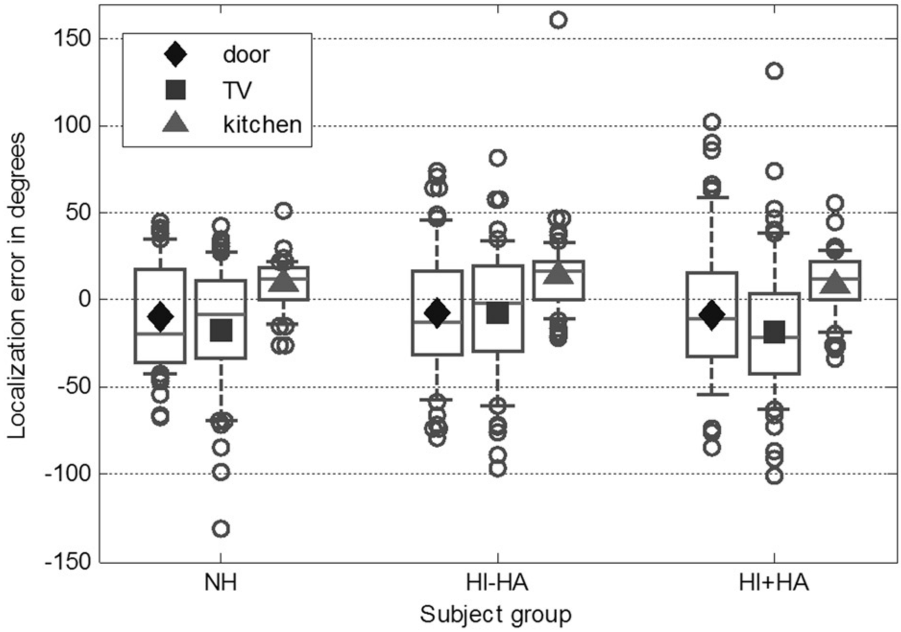

Figure 6. Box plot of localization errors of the phantom source presentation for the three subject groups NH, HI-HA, and HI+HA and three phantom source positions door (diamonds), TV (squares) and kitchen (triangles).

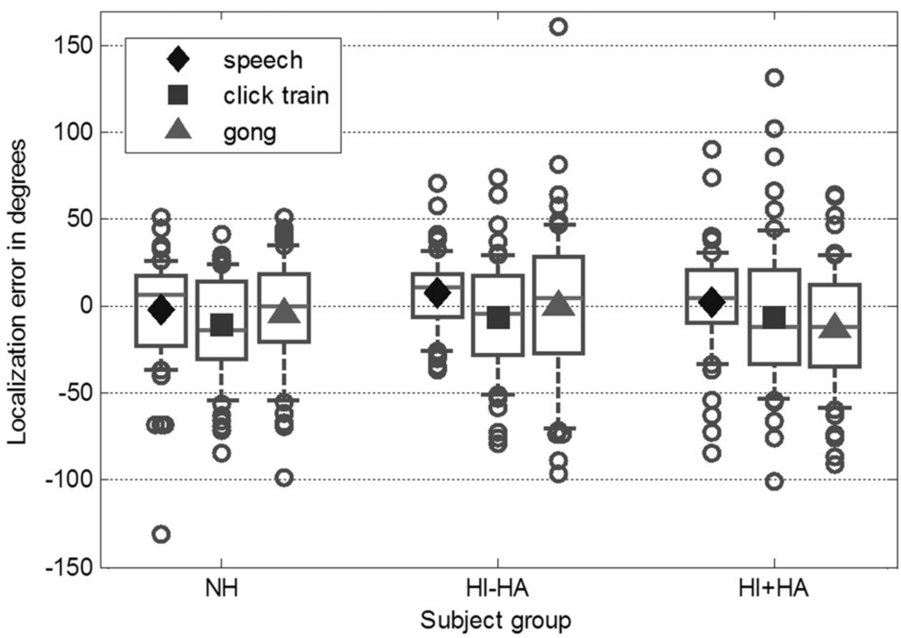

Figure 7. Box plot of localization errors of the phantom source presentation for the three subject groups NH, HI-HA, and HI+HA and three types of acoustic stimuli speech (diamonds), click train (squares) and gong (triangles).

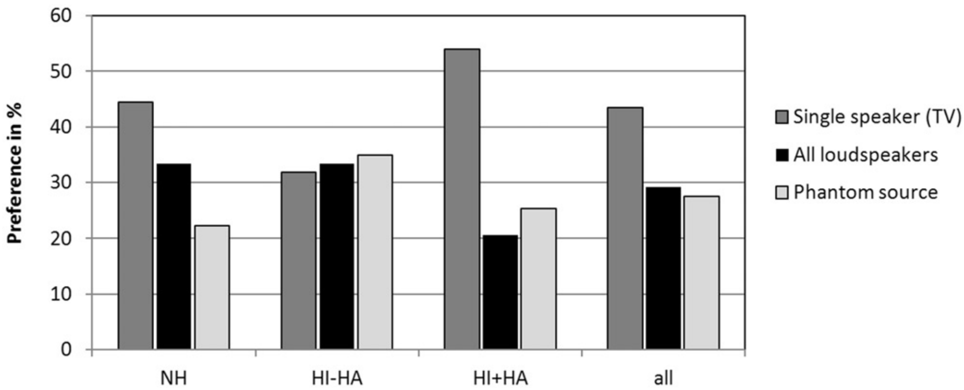

Figure 8. Preference percentages of the subject groups NH, HI-HA, HI+HA and all subjects for the three sound presentation concepts single speaker (TV), all speakers unisono and phantom source.

Preference determination

The results of the preference rating experiment are displayed in Figure 8. 43% of all subjects preferred the *single speaker* sound presentation, i.e. via TV speaker, 29% the *all unisono* concept and 28% the sound presentation via *phantom source* algorithm.

Broken down into the three subject groups, results show a clear top-down ranking of the *single-(TV) speaker*, *all unisono* and *phantom source* presentation concepts in the normal hearing group (NH), whereas unaided hearing impaired subjects (HI-HA) showed no preference for any of the concepts. Hearing impaired wearing hearing aids, however, preferred sound presentation via single TV loudspeaker about twice as much (54%) as presentations by all speakers or phantom source concept.

Acceptance assessment of the concepts

Regarding “situations with reference to locations at home” (e.g. door bell or alarm of a kitchen device), subjects generally gave the highest acceptance ratings for the *phantom source* concept, followed by the *all unisono* concept (s. Figure 9) with significantly lower ratings. The presentation via *single (TV) speaker* only was rated least acceptable for such situations by both hearing impaired groups, but not by normal hearing subjects. They rated both of the latter concepts basically equally.

Often named arguments for and against the respective concepts were, amongst others:

- Contra single (TV) speaker: “The TV won’t be running all the time.”
- Pro phantom source: “The phantom source tells you where to go.”
- Pro all speakers: “Using all speakers, one will be reminded anywhere in the apartment.”
- Contra all speakers: “All speakers playing would be too much.”

In the scenario *date reminder* as an example for a situation without reference to locations at home, all subjects gave significantly higher acceptance ratings for the concept of sound presentation by all loudspeakers than by

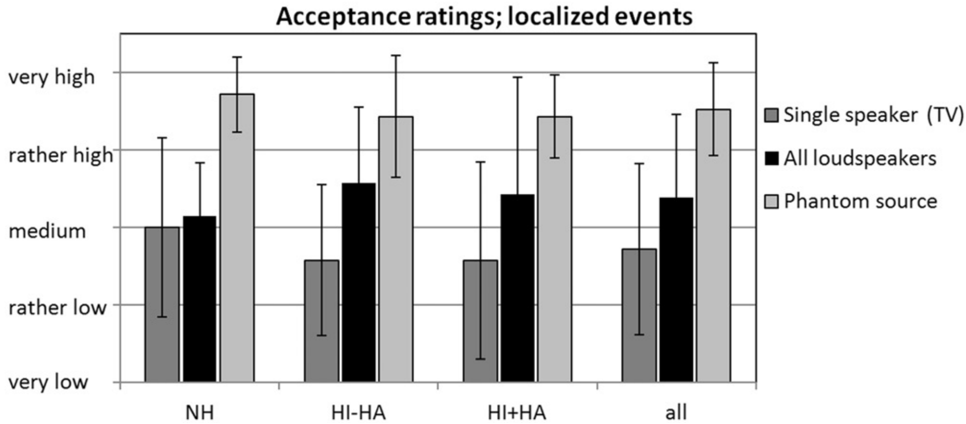

Figure 9. Average acceptance ratings for the three sound presentation concepts single (TV) speaker (mid grey), all loudspeakers (black) and phantom source (light grey) in the context of situations with reference to locations at home, obtained from the three subject groups normal hearing (NH, $N = 7$), unaided hearing impaired (HI-HA, $N = 7$), aided hearing impaired (HI+HA, $N = 7$) and all subjects (all, $N = 21$).

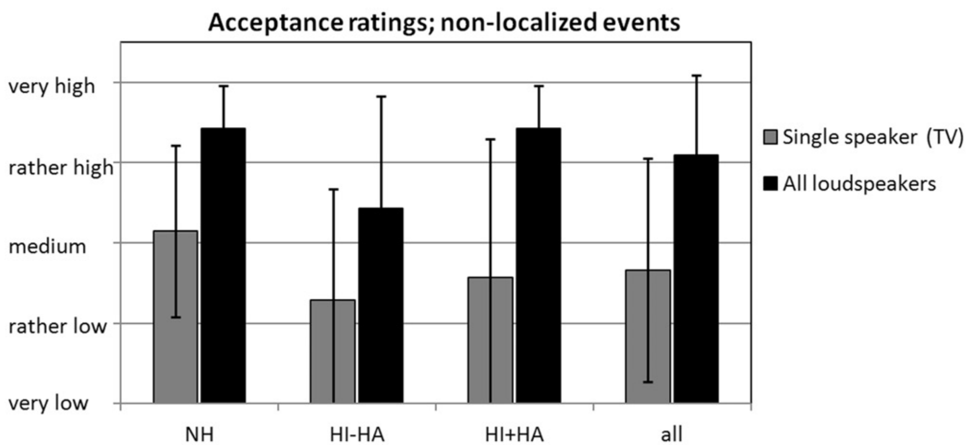

Figure 10. Similar to Figure 9, but in the context of situations without reference to locations at home.

a single (TV) speaker (cf. Figure 10). Often named reasons against the presentation by TV loudspeaker but for all speakers were:

- Contra single (TV) speaker: “The TV won’t be running all the time.”
- Contra single (TV) speaker: “I will not always be in the same room as the TV.”
- Pro all speakers: “Using all speakers, I will be reminded anywhere in the apartment.”

Preference for dynamic-compressed versus original signals

When comparing unprocessed audio signals with signals processed by multi-band dynamic-compression for hearing loss compensation, the majority

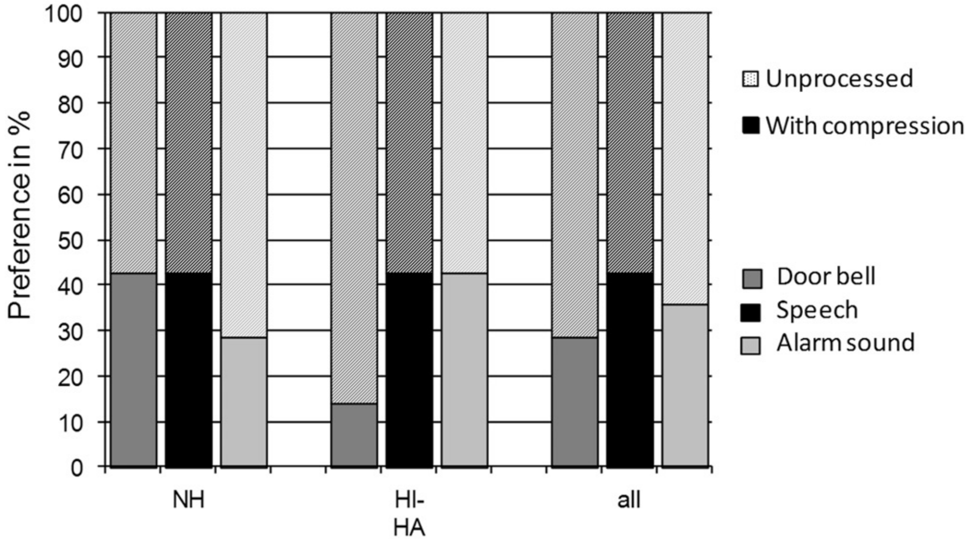

Figure 11. Preference percentages for dynamic-compressed signals (lower, filled parts of the bars) versus unprocessed signals (upper, dotted parts of the bars).

of subjects preferred the unprocessed versions (Figure 11). 43% of all subjects opted for the dynamic-compressed version of speech messages, 36% in case of alarm sounds and only 29% in case of the gong sound. Unprocessed speech was preferred equally by normal-hearing as well as hearing-impaired subjects, whereas unprocessed alarm sounds were especially preferred by normal-hearing and unprocessed gong sounds especially by unaided hearing-impaired subjects.

DISCUSSION

Localization experiment

The results of the localization experiments with the phantom sound source presentation concept have shown that listening position, direction of the phantom sound source as well as the type of the sound stimulus affect the localization performance. In contrast, hearing loss and wearing hearing aids or not does not have a significant influence on the localization accuracy. (Inter-group differences of absolute localization errors across all test situations did not reach 5% level of significance.) This finding, however, has to be considered with caution given the small sample size of the sub groups (7 subjects per group). Differences in localization accuracies between groups might reach level of significance as the sample size is increased. However, it was not a main aim of the present study to investigate possible differences between sub populations of potential users of the PAHA system. As stated earlier, the motivation to compose the overall group of 21 subjects with normal and typical (aided and unaided) hearing impaired subjects was rather to have a balanced, representative test sample. The obtained “pilot” results could be used as a basis to calculate required sample sizes for future studies that particularly aim at investigating the possible influence of hearing loss and/or wearing hearing aids with higher statistical power. If such studies should confirm the first results presented here, i.e. the absence of significant differences in localization accuracy between normal people and hearing aid users, this would be somewhat surprising, since especially hearing aid wearers might suffer from distortions of direction-indicating acoustical information (i.e. interaural level and time differences) by their hearing aid(s). The preservation of such directional acoustic cues has become (partly) possible rather recently by the latest generations of "true" binaural hearing aids, i.e. interacting bilateral hearing aids. In this study, however, the majority of the aided hearing impaired subject group HI+HA was not equipped with such hearing aids. On the other hand, the acoustical conditions of the environment were good (no interfering noise, low reverberation of the room) and the relative long duration of the acoustic test stimuli (ca. 4 s) allowed subjects to turn their heads towards the perceived direction of the sound source, which can help localizing sounds significantly.

In order to be able to localize a phantom sound source correctly, the algorithm has to "know" the exact listening position. In the presented experiments, this was the case for listening positions 1a and 2, whereas for listening position 1b, the algorithm parameters were not correctly set, but miss-adjusted by 50 cm. As expected, the subjects' localization accuracy was deteriorated at listening position 1b consequently; localization errors were even larger than those observed in the more "difficult" listening position 2 with very different distances to the loudspeakers. The results show that the presentation concept using a phantom source algorithm is sensitive to deviations between assumed and actual listening position. It can be concluded from the presented results that the accuracy of any automatic listener localization techniques has to better than 50 cm in order to allow for a sufficient localization accuracy of the phantom source. The exact tolerance range would have to be determined in a specific experiment.

The direction of the phantom sound source does also play a significant role. The experiments revealed a tendency that the perceived direction of the source seemed to be "dragged" towards the respective nearest apparent loudspeaker. Audio signals intended to originate from the kitchen, for example, were systematically localized more from the right, i.e. into the direction of the nearest loudspeaker (no. 1; cf. Figure 4), whereas the doorbell was localized systematically somewhat more to the left of the intended direction, i.e. also into the direction of the same loudspeaker. It can be assumed that because of the visibility of the loudspeaker and probable earlier experiences with conventional sound reproduction systems, it might have appeared more plausible to at least some of the subjects that the perceived sound originated from a physically real, visible sound source (here: a loudspeaker). Hence the observed results might be explained partly or even completely by audio-visual interaction. To exclude this assumed audio-visual interaction, loudspeakers could have been hidden with an opaque, but acoustically transparent material. However, this would have made the test environment less realistic and comparable to real home environments, where the loudspeaker positions will most probably be apparent (possibly somewhat less evident, though) or at least known by the user, too.

Another possible reason for systematic localization errors might have been that the sound level of the loudspeaker closest to the subject has actually been somewhat higher at the listener's position and hence "drag" the localization percept towards that loudspeaker's position, because, as pointed out in section "Concepts and algorithms for sound reproduction", an exact position-dependent loudspeaker level equalization (which the phantom source algorithm relies on) can be difficult in real environments.

The type of the audio signal had a significant effect on the localization accuracy as well. Speech was associated with the lowest localization errors, possibly because human listeners are highly trained to the perception and localization of speech in everyday life. Click trains were localized significantly poorer, although the broadband, transient character of this artificial stimulus should actually be beneficial for localization, which is the reason why this type of stimulus is commonly used in psychoacoustic localization experiments (21). Maybe the subjects' familiarity with speech outweighed the theoretical psychoacoustical advantages of the very artificial, unfamiliar click stimuli in the present study. The localization of the gong sound posed a big problem especially for unaided hearing impaired subjects; the reported perceived directions varied highly. In contrast, the localizations of this stimulus were the most consistent and accurate ones in the normal hearing group. Whether there is an actual systematic correlation between localization accuracy and hearing loss for this type of stimulus remains yet unclear. A possible worse audibility of the gong stimulus for the mostly mild-to-moderately impaired subjects does not appear likely, since the frequency spectrum of the gong signal did not differ significantly from that of speech, regarding the relative power portion at higher frequencies.

Finally, it should be noted that the general localization accuracy in real life practice might be somewhat lower than observed here due to a possible psychological effect of the experimental condition on the performance of the subjects, who, being watched, might be especially motivated to perform as good as possible.

Preference and acceptance of the presentation concepts

The main purpose of the presented study was to determine which of the described sound presentation alternatives would be most suitable and preferred by potential users of the PAHA system. Background and purpose of the study were explained to all subjects prior to the experiments. Subjects were asked to imagine themselves as actual users of the PAHA system when assessing their acceptances and preferences for the respective methods.

Reported preferences for the phantom source concept and for presentation via all loudspeakers *unisono* were similar. The TV (single-loudspeaker concept) was the most preferred presentation option, maybe because it was the closest of all sources in two of the three tested listening positions and therefore possibly the one with the best reception. (Sound levels of all loudspeakers at the listening position were about the same due to the prior level equalization, though, but the direct-to-reverberant energy ratio (given by the room acoustics) is generally increases and hence improves intelligibility and clearness, as the distance between sound source and listener decreases.)

When taking a look at subjects' general acceptances for the different presentation concepts from a more theoretical, conceptual view point, independent of the concrete sample sounds that were demonstrated to the subjects in the prior preference rating experiment, a completely different picture is found. Here, the phantom source concept providing a kind of acoustical “guidepost” for situations with reference to locations at home (e.g. door bell or alarm of a kitchen device) got the highest acceptance ratings, i.e. was preferred over the other concepts, whereas the single TV loudspeaker was rated the least. When asked for reasons, some subjects stated that the audio signal emitted from a phantom source has a directional component and thus conveys additional information about the location of the event (e.g. warning from the kitchen), whereas presentations via TV loudspeaker would depend on a running TV. However, this objection was due to the misconception that the TV would have to be running if sounds are be played by its loudspeaker, whereas all other loudspeaker would generally be “active” all the time anyway. However, in a later realization of the system prototype tested here, this does not necessarily have to be the case. Another stated, more valid objection was the expectation of many subjects that messages played by the phantom source would be noticeable everywhere in the apartment (provided that the whole apartment was equipped with loudspeakers), whereas the TV loudspeaker, being a source at a fixed location, would only be audible well in the same room.

For the purpose of acoustical notifications of events *without* reference to locations at home (e.g. reminder of an appointment), the presentation via all loudspeakers got the highest acceptance ratings, because subjects expected that they could hear the notification anywhere in the apartment this way, independent of the TV running if they were in the same room as the TV or not. Again, there was the misconception of a permanently running TV being a prerequisite for notifications via the TV loudspeaker.

It should be noted that the phantom source presentation method was not described to the subjects or even emphasized as a “special”, new method. Also in view of the lower preference for this method observed in the direct comparison experiment for preference determination, we do not think that the obtained evaluation results were biased towards higher subjective ratings of the phantom source method.

Preference of unprocessed versus dynamic-compressed audio signals

Overall, subjects preferred sound presentations without dynamic compression for the compensation of hearing losses. Some subjects stated that dynamic-compressed speech was more intelligible, but had a less comfortable sound quality. This observation of a rather low “spontaneous acceptance” of new, unfamiliar sound is in line with the typical experience made with first-time hearing aid users, who most often appreciate the better audibility of soft sounds thanks to the amplification of the hearing aid, but rather prefer the “original” sound quality they are accustomed to. However, compensation of hearing losses that typically increase with frequency will alter the sound quality; the sound becomes brighter, sharper. This is often getting to the point where first-time hearing aid users reject complete hearing loss compensation at high frequencies, even if this would allow for better speech understanding, because they regard the amplified sound as being too sharp. As a consequence, the process of fitting a hearing aid to an individual hearing loss is typically done in several steps over an acclimatization period of several weeks to months, in which the hearing aid user can get used to the new sound quality, while the amplification settings of the hearing aid are increased slowly, step by step. The group of subjects with age-appropriate normal hearing (i.e. only mildly elevated hearing thresholds mainly at higher frequencies) of the presented study might have hardly benefited from the dynamic compression, but possibly only been irritated by the altered sound quality. Some subjects of the unaided hearing impaired subjects (HI-HA), however, seemed to have benefited by the signal processing and reported a better speech understanding. In addition to the stated negative effect of an altered sound quality, hearing losses of the unaided hearing impaired subject group varied markedly between subjects, while the dynamic compression setting was fitted only to the average hearing loss per subject group to limit the overall effort. This simplification might have lead to sub-optimal fittings for some individual subjects and consequently to lower acceptances by these subjects for the processed sounds.

What might impede the acclimatization to dynamic-compressed audio signals in the context of the GAL concept is the circumstance that in the user's environment only the sound presentations of the GAL system would be adapted to the user's hearing loss, in contrast to all other sounds. On the other hand, this could be a chance to demonstrate the benefits of assistive listening technologies to the users, which might lower the barrier for hearing impaired persons to decide to get aided in general by using hearing aids. This could help tackling the still existing problem of a much too low hearing aid provision rate: In Germany, only approximately 25% of all hearing impaired persons that would benefit from hearing aids are actually aided.

Higher acceptances for dynamic-compressed audio can be expected in case of individualized fittings and after a period of acclimatization. The actual benefit of the tested system for hearing impaired users could and should then be evaluated by means of formal speech intelligibility measurements and assessments of, e.g. the perceived listening effort.

SUMMARY AND CONCLUSIONS

- Three different sound presentation concepts as part of an automated personal activity and household assistant for ambient assisted living have been developed and evaluated in a user study with elderly normal-hearing and hearing-impaired subjects.
- The localizability of the phantom source was found to be good, as long as the exact listeners' position is known and adequate acoustic stimuli (natural, familiar, e.g. speech) are used. Errors in the determination of the listener's position of 0.5 m (possibly even smaller) lead to significant larger errors and difficulties in sound source localization by the users. No systematic differences in localization performances were found between aided and unaided (hearing impaired and normal hearing) subjects.
- When comparing acoustic demonstrations of the presentation concepts directly, most of the test subjects preferred sound presentation via TV loudspeaker. However, regarding the general concept, the highest acceptance ratings were given for phantom sources when notifications of events with reference to locations at home were concerned. Regarding events

without reference to locations, sound playback by all loudspeakers *unisono* was accepted the most.

- Audio signals that were pre-processed by multi-band dynamic compression to compensate for elevated hearing thresholds were mostly rejected in direct A/B comparisons with unprocessed signals, although processed speech was often described as being clearer.
- One centrally located loudspeaker should be installed per room, ideally in an ambient way. Where available, the loudspeaker should be placed close to or behind the TV or other visual user interface of the system, respectively, in order to ensure a coherent appearance of the overall system. This setup would have the additional advantage to be cost-efficient and easily changeable.
- The compensation of possible hearing losses by signal pre-processing should be adapted and fine-tuned to the individual user. Benefit and acceptance of such pre-processing should be investigated further in a more specific, detailed study.

ACKNOWLEDGEMENTS

The authors wish to thank the two anonymous reviewers for their helpful comments on earlier versions of this work. The Lower Saxony research network "Design of Environments for Ageing" acknowledges the support of the Lower Saxony Ministry of Science and Culture through the "Niedersächsisches Vorab" grant programme (grant ZN 2701).

Local ethical approval was granted for the study.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES

1. Boll S, Heuten W, Meyer EM, Meis M. Development of a multimodal reminder system for older persons in their residential home. Inform Health Social Care 2010; 35:104–24.
2. Goetze S, Moritz N, Appell JE, et al. Acoustic user interfaces for ambient-assisted living technologies. Inform Health Social Care 2010;35:125–43.
3. Merchel S, Groth S. Adaptively adjusting the stereophonic sweet spot to the listener's position. J Audio Eng Soc 2010;58:809–17.
4. Pulkki V. Virtual sound source positioning using vector base amplitude panning. J Audio Eng Soc 1997;45:456–66.
5. Van Leest AJ. On amplitude panning and asymmetric loudspeaker set-ups. Presented at the 119th Convention of the Audio Engineering Society; 2005 October 7–10; New York, NY, USA. Convention Paper 6613.
6. Gerzon MA. Periphony: with-height sound reproduction. J Audio Eng Soc 1973;21: 2–10.
7. Fellgett P. Ambisonics. Part one: general system description. Studio Sound 1975;1: 20–2.

8. Spors S, Rabenstein R, Ahrens J. The theory of wave field synthesis revisited. 124th Convention of the Audio Engineering Society; 2008 May 17–20, Amsterdam.
9. Stackelberg Hv. Hörtest 1985. Bonn: Deutsches Grünes Kreuz; 1986.
10. Davis AC. The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. Int J Epidemiol 1989;18:911–17.
11. Hougaard S, Ruf S, Egger C. EuroTrak + JapanTrak 2012 – world's largest multi-country consumer survey about hearing, hearing loss and hearing aids. Hear Rev USA 2013;20:16–26.
12. Gerzon MA. The design of distance panpots. Presented at the 92nd Convention of the Audio Engineering Society. 1992 March 24–27; Vienna. Preprint 3308.
13. Theile G, Plenge G. Localization of lateral phantom sources. J Audio Eng Soc 1977; 25:196–200.
14. Grimm G, Herzke T, Berg D, Hohmann V. The master hearing aid: a PC-based platform for algorithm development and evaluation. Acta Acustica 2006;92:618–28.
15. Byrne D, Dillon H, Ching T, et al. The NAL-NL1 procedure for fitting non-linear hearing aids: characteristics and comparisons with other procedures. J Am Acad Audiol 2001;12:37–51.
16. World Health Organization. Grades of hearing impairment. Hearing Networks News 2001;1:144–57.
17. Lewald J, Ehrenstein WH. Auditory-visual spatial integration: a new psychophysical approach using laser pointing to acoustic targets. J Acoust Soc Am 1998;104: 1586–97.
18. Seeber B. A new method for localization studies. Acta Acustica 2002;88:446–50.
19. Backhaus K, Erichson B, Plinke W. Multivariate analysemethoden. Berlin: Springer; 2003.
20. Bortz J. Statistik für Human- und Sozialwissenschaftler. 6th ed. Berlin: Springer; 2005.
21. Otten J. Factors influencing acoustical localization. [PhD thesis]. Oldenburg: University of Oldenburg; 2001. Available from: http://oops.uni-oldenburg.de/335/ [last accessed 2013 September 30].