The Effect of Hearing Loss on Localization of Amplitude-Panned and Physical Sources

Gregory M. Ellis, PhD1  Pamela E. Souza, PhD, CCC-A1

1Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois

Address for correspondence Gregory M. Ellis, PhD, gregory.ellis@northwestern.edu


Abstract

Background Clinics are increasingly turning toward using virtual environments to demonstrate and validate hearing aid fittings in “realistic” listening situations before the patient leaves the clinic. One of the most cost-effective and straightforward ways to create such an environment is through the use of a small speaker array and amplitude panning. Amplitude panning is a signal playback method used to change the perceived location of a source by changing the level of two or more loudspeakers. The perceptual consequences (i.e., perceived source width and location) of amplitude panning have been well-documented for listeners with normal hearing but not for listeners with hearing impairment.

Purpose The purpose of this study was to examine the perceptual consequences of amplitude panning for listeners with hearing statuses from normal hearing through moderate sensorineural hearing losses.

Research Design Listeners performed a localization task. Sound sources were broadband 4 Hz amplitude-modulated white noise bursts. Thirty-nine sources (14 physical) were produced by either physical loudspeakers or via amplitude panning. Listeners completed a training block of 39 trials (one for each source) before completing three test blocks of 39 trials each. Source production method was randomized within block.

Study Sample Twenty-seven adult listeners (mean age 52.79, standard deviation 27.36, 10 males, 17 females) with hearing ranging from within normal limits to moderate bilateral sensorineural hearing loss participated in the study. Listeners were recruited from a laboratory database of listeners that consented to being informed about available studies.

Data Collection and Analysis Listeners indicated the perceived source location via touch screen. Outcome variables were azimuth error, elevation error, and total angular error (Euclidean distance in degrees between perceived and correct location). Listeners’ pure-tone averages (PTAs) were calculated and used in mixed-effects models along with source type and the interaction between source type and PTA as predictors. Subject was included as a random variable.

Results Significant interactions between PTA and source production method were observed for total and elevation errors. Listeners with higher PTAs (i.e., worse hearing) did not localize physical and panned sources differently whereas listeners with lower
Among research and clinical hearing professionals, there is increasing interest in measuring communication and treatment outcomes under “real-life” conditions. One approach attempts to quantify aspects of the listener’s actual listening environment. This category includes body-worn devices that capture acoustic characteristics of the listener’s environment and ecological momentary assessment in which the listener is periodically prompted to answer questions about their listening experience. In a different approach, researchers have attempted to create ecologically relevant environments within a laboratory space. Some in-laboratory virtual environments likely fall short of a true everyday environment because of equipment or space constraints. Many such virtual environments require large numbers of speakers to accurately represent sources. Sound booths that are large enough to house such a setup are expensive to acquire and can be difficult to install in buildings that do not already have such facilities. For these reasons, setups requiring dozens of loudspeakers are limited to resource-rich laboratories and minimally accessible to clinics.

Drawing partly from innovations in virtual reality (see, for example, Frank et al), there is keen interest in creating semi-virtual environments that require less hardware than traditional in-laboratory “realistic” scenarios. Many earphone situations have been devised, but those are unsuitable for testing benefit of devices that must be worn in soundfield (e.g., hearing aids, cochlear implants, assistive devices). A promising approach for this kind of research is to use a combination of physical loudspeakers and virtual acoustics to build the desired space. Using this technique, there is potential for high flexibility with minimal hardware and cost. Relatively speaking, it is easy to construct a virtual environment using a set of virtual sources through amplitude-panning techniques like vector-based amplitude panning and manifold-interface amplitude panning.

When matching an adjustable virtual source to a fixed physical source, listeners with normal hearing make relatively small localization errors across a wide range of loudspeaker arrangements; however, effects of loudspeaker arrangement have been observed. Localization errors tend to get larger as the speaker pair or triplet is moved laterally off the sagittal plane (i.e., nearer to the listener’s right or left; +90 or −90 degrees, respectively) or as the loudspeaker pair is separated by a larger angle. This effect of loudspeaker separation on change in localization errors is most noticeable for amplitude panning in elevation, especially when the loudspeakers are far apart. Research on amplitude panning generally shows that localization errors are smaller for broadband signals (i.e., speech, noise, etc.) than for narrowband signals (i.e., single tones, bandpass noise, etc.). Research on listeners with normal hearing shows promise for building a virtual space out of a combination of physical and panned sources; however, the perceptual effects of amplitude panning-based virtual sources on listeners with impaired hearing are poorly understood.

It is important to understand if listeners with hearing impairment show the same pattern of localization errors, especially if hearing clinics are going to use this technology to generate virtual environments for testing and validating hearing aid fittings. Currently, data which examine the consequences of amplitude panning for listeners with hearing loss are lacking. The present study seeks to better understand the localization of amplitude panned virtual sources among listeners with a range of hearing thresholds.

**Methods**

**Participants**

Twenty-seven adults (mean age 52.79, standard deviation 27.36, 10 males, 17 females) participated in the study. Most of the listeners had symmetric hearing (defined as between-ear pure-tone averages [PTAs] 0.5–1, and 2 kHz) and individual test frequency differences ≤15 dB up to 8 kHz. There was a single listener with larger asymmetries in two pure-tone frequencies (20 dB at 2 kHz and 30 dB at 3 kHz) (see the “Analysis” section for data considerations). PTAs ranged from normal hearing (1.67 dB) to moderate loss (48.33 dB). PTA was significantly correlated with age (r = 0.72, p < 0.001). Individual audiograms are plotted in –Fig. 1. All listeners spoke English as their primary language, completed an informed consent process approved by the Northwestern University Institutional Review Board, and were compensated $15/hour for their time. All study participation was completed in one session that took no more than an hour.

**Procedures**

**Stimuli**

Localization stimuli were 1-second long, amplitude-modulated broadband white noise bursts. Broadband noise bursts were used because previous studies have used similar stimuli and because broadband sources are easier to localize than narrowband stimuli, thereby reducing test-retest variability and improving understanding of the task for the
listener.\textsuperscript{13–15,17,20–22} Stimuli were 4 Hz amplitude-modulated with 100% modulation depth in sine phase. Four-hertz amplitude-modulation and 100% depth were used to give listeners multiple “looks” at each stimulus. Sine phase was used so that the stimuli ramped on and off smoothly.

Stimuli were presented from a loudspeaker array. Listeners were seated in the middle of a large (4.9 m × 4.3 m × 2.7 m) sound-attenuated room with 37 loudspeakers. Six loudspeakers and one subwoofer were located on each wall, with nine loudspeakers located in the ceiling of the room. Sound sources were a mixture of real and virtual sources.

Virtual sources were generated using SpaceMap, a spatialization tool within Meyer Sound’s CueStation. SpaceMap allows an experimenter to create virtual environments. An open-source version expanding on SpaceMap’s functionality is available for free to the interested reader.\textsuperscript{12} To create virtual sounds, first the physical loudspeaker locations must be provided to SpaceMap. Next, the experimenter generates a loudspeaker triplet within SpaceMap by choosing a set of three physical loudspeakers. A virtual sound source can be placed anywhere within—or on the sides of—a triangle formed by the three loudspeakers. The experimenter then specifies the desired location of the sound source within the boundaries of the triangle. Based on this location, SpaceMap calculates the output levels of the three loudspeakers to place the source at the specified location.

In total, 39 source locations (14 real and 25 virtual) were tested. Sound sources spanned a grid from −90 to +90 degrees azimuth in 15-degree steps and −20 to +20 degrees elevation (below ear-level and above ear-level, respectively) in 20-degree steps (\textsuperscript{Fig. 2}). This span of sources was chosen to avoid errors due to front-back confusions.

**Localization Task**

Listeners faced the front of the room and were allowed to freely move their head during stimulus presentation. Although the head is traditionally fixed in localization experiments,\textsuperscript{23,24} the literature on localizing virtual sources suggests that allowing for free head movement strengthens the percept of the virtual source.\textsuperscript{16,17} It is also the expected scenario for use of this technique in a clinical context.\textsuperscript{25–28} For these reasons, free head movement was deemed to be an acceptable experimental choice in the present study.

Listeners performed the task in four blocks. There were 39 trials in each block. Each source location was presented once per block in a random order. Each block took less than 10 minutes to complete. Listeners were told to listen for the location of a sound in the room in front of them. Listeners reported the perceived location of the sound on a touchscreen (Samsung Galaxy Tab E) using a GUI developed in MATLAB (\textsuperscript{Fig. 3}). All listeners completed a training block of 39 trials with feedback to familiarize them with the response method. The experimenter sat with the listeners and guided them in the task during training. After training, listeners were told that they would no longer receive feedback and continued at their own pace through the remaining three experimental blocks. Listeners were offered a break between blocks.

**Analysis**

Three types of localization error were used in the analysis: (1) degree error in azimuth (azimuth error), the error in degrees along the horizontal plane (left/right), (2) degree error in elevation (elevation error), the error in degrees along the sagittal plane (up/down), and (3) overall degree error (total error), the combined error in degrees between the source and the response. Here, total error is calculated as the square root of the sum of squared azimuth error and squared elevation error. Suppose, for example, the location of the source is 0 degrees elevation and +15 degrees azimuth, and the listener responds at +15 degrees elevation and +35 degrees azimuth. In that case, azimuth error is +35 degrees (listener’s response) minus +15 degrees (source location) for a +20-degree error. Elevation error is +15 degrees (listener’s response) minus 0 degrees (source location) for a +15-degree error. For analysis, the signs on azimuth and elevation errors were dropped to track only deviation from the source rather than deviation and direction. Total error is calculated as a line representing the combined azimuth and elevation error. In this example, total error would be $\sqrt{(20^2 + 15^2)} = 25^\circ$. 

![Fig. 1](Individual audiograms. Left ear audiometric thresholds plotted on the left with "x's." Right ear audiometric thresholds plotted on the right with circles.)
Mean performance change (panned – physical) for each individual in the three error domains (total, azimuth, and elevation) is plotted below ordered by PTA (►Fig. 4). Bars to the right of the zero line indicate larger errors in the panned listening conditions than in the physical listening conditions. The listener with some asymmetry mentioned in the participants section has a PTA of 29.2 dB. Visual inspection of their performance indicates that they were not likely to skew the analysis in any meaningful way, so they were included in all analyses. Many of the listeners with lower PTAs show larger errors for sources generated with amplitude panning than with physical sources. Listeners with higher PTAs, however, do not show as consistent a pattern of results. Many such individuals with higher PTAs show no difference in localization accuracy between physical and panned sources. This is most noticeable when responses are represented as total error (rightmost panel of ►Fig. 4).

Mixed-effects models were used to analyze the effects of hearing status and source type (physical/panned) on the three types of localization errors. Hearing status was defined...
was the random effect. had the same with each type of error as a dependent variable. All models term will be referred to as "panning and physical sources, respectively. This model "or a "structural threshold of 5, —— widely accepted "criterion for each error type (azimuth, elevation, and total) sorted by increasing pure-tone average (PTA). Change in error is plotted in each column. Bars to the right of the zero line indicate that the listener is more accurate in the physical listening conditions than the panned conditions. Bars to the left of the zero line indicate the opposite. Note that the bars tend to get smaller as PTA increases, indicating listeners with hearing loss have more difficulty distinguishing between physical and panned sources when asked to identify the location of the sound source via three-frequency PTA. Sources were coded as either a "0" or a "1" to represent sources generated with amplitude panning and physical sources, respectively. This model term will be referred to as "physicality." To look at the different types of errors, three separate models were run with each type of error as a dependent variable. All models had the same fixed and random effects. Fixed effects were PTA, source verity and the interaction between them. Subject was the random effect.

**Results**

Since age and PTA were significantly correlated (r = 0.72, p < 0.001), several steps were taken to examine whether differences in age contributed to performance differences. First, the variance inflation factor (VIF) was examined to determine whether including age and PTA in the same model would be statistically problematic. It was found that in models using physicality, age, PTA, and subject, VIF was low for all four factors: 1.00, 2.12, 2.46, and 1.51, respectively. Because none of these exceeded the older—though still widely accepted—threshold of 10, the newer, more conservative threshold of 5, all terms were deemed accept-

able to keep in the model. Next, likelihood ratio tests were conducted to determine whether adding age significantly improved any of the three models. Adding age did not significantly improve any of the models and was therefore excluded from all further analyses (azimuth error model: chi-square (6) = 3.46, p = 0.749, ΔR² = 0.0003; elevation error model: chi-square (6) = 9.81, p = 0.133, ΔR² = 0.0007; total error model: chi-square (6) = 1.44, p = 0.963, ΔR² = 0.0003).

Next, likelihood ratio tests were conducted to determine whether the subject random effect significantly improved the models. Adding subject as a random variable significantly improved the models predicting azimuth error (chi-square (10) = 287.88, p < 0.001, ΔR² = 0.094), elevation error (chi-square (10) = 120.66, p < 0.001, ΔR² = 0.046), and total error (chi-square (10) = 331.29, p < 0.001, ΔR² = 0.1010). Subject was included in all subsequent models.

The mixed-effects model predicting azimuth error explained significantly more variance than a constant model (chi-square (13) = 614.75, p < 0.001, ΔR² = 0.1926). The effects of PTA and physicality are plotted in - Fig. 3. The figure shows that, across PTA, listeners on average make larger azimuth errors when localizing panned sources (gray symbols and lines) than when localizing physical sources (black symbols and lines). These errors increase as PTA increases. The mixed-effects model predicting elevation error also explained significantly more variance than a constant model (chi-square (13) = 405.79, p < 0.001, ΔR² = 0.1326; - Fig. 6). - Fig. 6 shows a slightly more complicated relationship between PTA and source verity when predicting elevation errors. The interaction between PTA and source verity is significant (b = 0.28, t(3150) = 5.6062, p < 0.001) and positive, indicating statistically that errors grow more quickly for physical sources than panned sources. The interaction can be seen in - Fig. 6: elevation errors increase as PTA increases regardless of source verity; however, the slope is greater for elevation errors made when localizing physical sources than the slope for panned sources.

![Fig. 4](image)

**Fig. 4** Individual mean change in performance (Δ = panned – physical) for each error type (azimuth, elevation, and total) sorted by increasing pure-tone average (PTA). Change in error is plotted in each column. Bars to the right of the zero line indicate that the listener is more accurate in the physical listening conditions than the panned conditions. Bars to the left of the zero line indicate the opposite. Note that the bars tend to get smaller as PTA increases, indicating listeners with hearing loss have more difficulty distinguishing between physical and panned sources when asked to identify the location of the sound source.

![Fig. 5](image)

**Fig. 5** Effect of pure-tone average (PTA) and amplitude panning on azimuth errors in sound source localization. Upward-facing triangles represent the estimated means for the panned sources. Downward-facing arrows represent the estimated means for the physical sources. Dotted lines are lines of best fit for the physical (black) and panned (gray) sources. For azimuth localization, there was no observed interaction between PTA and source verity indicating that, in general, listeners are sensitive to source verity in the horizontal plane.
Fig. 6  Effect of pure-tone average (PTA) and amplitude panning on elevation errors in sound source localization. Key is the same as for the previous figure. For elevation localization, there was an interaction between PTA and source verity.

Fig. 7  Effect of pure-tone average (PTA) and amplitude panning on total error in sound source localization. Key is the same as for the previous figure. Note that the listeners with lower PTAs generally see larger errors for panned sources than physical sources; however, this trend does not hold for listeners with higher PTAs (over ~20 dB).

The mixed-effects model predicting total error explained significantly more variance than a constant model (chi-square (13) = 816.78, p < 0.001, ΔR² = 0.2436; Fig. 7). Another significant interaction is seen in Fig. 7 (b = 0.23, t (3150) = 4.0478, p < 0.001). The relationship between PTA, source verity, and total error is similar to that between PTA, source verity, and elevation error. The coefficients for all three models are summarized in Table 1.

To put these coefficients in a larger context, this statistical model predicts that a listener localizing a panned source with a 50 dB hearing level (HL) PTA will, on average, make localization errors 13 degrees larger (confidence interval [CI]95 = [8.4, 17.6]) than a listener with a 10 dB HL PTA. For a real source, this error increases to 21 degrees (CI95 = [13.2, 31.2]). These numbers are calculated by using the following formula:

\[
\text{Response Error} = b_0 + b_1 \times \text{PTA} + b_2 \times \text{Source Verity} + b_3 \times \text{Physicality} + b_4 \times \text{PTA} \times \text{Physicality},
\]

where \(b\) indicates the appropriate coefficient from Table 1 and \(X\) indicates a value for physicality (0 or 1 representing panned and virtual sources, respectively) or a value for PTA.

The fact that the slope is larger for physical than virtual sources is driven by the fact that the listeners with lower PTAs (i.e., better hearing) generally show differences between physical and amplitude-panned source localization whereas the listeners with higher PTAs (i.e., worse hearing) do not. This can be seen in the figures above by the nonoverlapping error bars for the listeners with lower PTAs and the almost exclusively overlapping error bars for the listeners with higher PTAs. It is worth noting as well that listeners with higher PTAs are generally more variable than those with lower PTAs. Increased variability among listeners with higher PTAs is also evident in previous localization studies.

**Discussion**

The present study observed that listeners with better hearing are more likely to have larger localization errors for virtual sources generated with amplitude panning than physical loudspeakers. This result differs from previous research on amplitude panning demonstrating small or no noticeable differences in localization between sources generated with amplitude panning and physical loudspeakers; however, there is a methodological difference between the present study and past research. Here, listeners localized sources after hearing the sound once; similar to the situation that might occur in everyday listening when a listener orients to a transient sound. Previous studies used adjustment, where listeners moved virtual sources to coincide with a physical source. Adjusting source location will produce smaller errors because listeners can continuously monitor the locations of both the physical and virtual sources.

The present study demonstrates that listeners with higher audiometric thresholds (i.e., worse hearing) are likely to localize physical and panned sources with the same degree of error. The statistical model presented above can be compared with previous work examining localization ability in listeners with hearing impairment by plugging in the reported sample’s PTA and whether the sources were real or not (physicality). Häusler et al report their listeners to have:

**Table 1** Results of the three models predicting errors in localization as a function of PTA and amplitude panning

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Physicality</th>
<th>PTA</th>
<th>Physicality × PTA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Azimuth error</strong></td>
<td>(b = 8.90 ) [6.98, 10.81]</td>
<td>(b = -3.80 ) [−5.92, −1.89]</td>
<td>(b = 0.25 ) [0.13, 0.37]</td>
<td>(b = 0.03 ) [−0.05, 0.11]</td>
</tr>
<tr>
<td><strong>Elevation error</strong></td>
<td>(b = 10.74 ) [9.29, 12.18]</td>
<td>(b = -4.26 ) [−6.70, −1.82]</td>
<td>(b = 0.14 ) [0.08, 0.20]</td>
<td>(b = 0.28 ) [0.18, 0.38]</td>
</tr>
<tr>
<td><strong>Total error</strong></td>
<td>(b = 15.62 ) [13.79, 17.54]</td>
<td>(b = -5.56 ) [−8.58, −2.54]</td>
<td>(b = 0.33 ) [0.21, 0.44]</td>
<td>(b = 0.23 ) [0.12, 0.34]</td>
</tr>
</tbody>
</table>

Abbreviation: PTA, pure-tone average.

Note: Bolded values are significant at the \(p < 0.001\) level. Bracketed numbers are 95% confidence intervals for the coefficient estimates.
have a PTA of 35 dB HL. Their listeners located physical sources. For these parameters, the statistical model presented here predicts that listeners should make azimuth errors of 13.85 degrees on average (CI95 = [5.61, 21.87]). Häusler et al.\(^2\) found that their listeners had an azimuth minimum audible angle (MAA) ranging from 5 degrees to over 30 degrees depending on source location. For elevation, the statistical model predicts that listeners should make errors of 21.18 degrees on average (CI95 = [11.69, 30.66]). Häusler et al.\(^2\) found that elevation MAAs cover a range from 1 to 30 degrees. Thus, the statistical model’s predictions align nicely with Häusler et al’s work.

The data presented here likely extend to speech localization. Previous studies have shown that broadband signals are easier to localize than narrowband signals,\(^3\)–\(^7\),\(^15\)–\(^17\),\(^20\)–\(^22\) especially if those signals have energy in the relevant frequency ranges for interaural timing and interaural level differences—below approximately 1.5 kHz and above approximately 3 kHz, respectively.\(^33\) Both broadband noise (studied here) and speech have energy covering this range. Speech localization is of particular interest to clinical and translational work using virtual panning, given the relevance of speech localization to communication in realistic environments with multiple talkers. Such work presents an interesting direction for future study.

The present study has a few limitations. First, front-back confusions could have occurred and not been recorded, though they likely occurred infrequently during the experiment. There are several reasons to think that the incidence rate of front-back confusions was low. First, listeners were explicitly told that all sources would come from the front and the block of training trials was designed to provide listeners with sufficient training to recognize that all source locations were in the front hemifield. Listeners were also allowed to move their head freely during the experiment. Free head movement has been shown to help resolve front-back confusions, further reducing the likelihood of their occurrence in this case. Future studies examining localization of virtual sources may find it useful to note when listeners have a front-back confusion so these trials can be omitted or analyzed appropriately. Second, listeners in the present study were not using assistive devices (e.g., hearing aids, cochlear implants, remote microphones, etc.). It is highly likely that using these devices would change these listeners’ localization accuracy, either through deliberate directional processing or acoustic effects of earpiece coupling. Given that current-generation devices offer a large variety of different directional effects, evaluation of such factors was beyond the scope of the present study.

In summary, the results of this study provide an initial evaluation of the effect of virtual panned sources for listeners with hearing loss. As hearing impairment becomes more severe, listeners localize panned and physical sources with similar error magnitudes, albeit in quiet and with a single source. The finding that the errors in localizing panned and physical sources are not different for listeners with mild-to-moderate hearing losses suggests that amplitude panning could be an appropriate method for constructing virtual environments for these listeners. However, multiple sources or noisy environments may affect the listener’s ability to localize sound, either through acoustic interference or the need to employ cognitive processing to direct or inhibit attention. More work is needed to extend these initial results to noisy or complex environments or when the listeners are wearing assistive devices.

Conclusion

When performing a localization task, listeners with lower PTAs show significantly larger errors for virtual sources generated with amplitude panning than for physical sources. Listeners with higher PTAs show no significant difference in errors between virtual sources generated with amplitude panning and physical sources. This lack of error difference could serve as a justification that supports recent interest in creating within-clinic virtual environments that are based on amplitude panning.

Funding

This work was supported by funds from the NIDCD, grant R01: DC006014.

Conflict of Interest

Dr. Ellis reports grants from National Institute on Deafness and other Communication Disorders, during the conduct of the study. Dr. Souza reports grants from National Institutes of Health, during the conduct of the study; personal fees from Phonak, outside the submitted work.

Acknowledgments

The authors would like to thank the members of the Hearing Aid Laboratory for their insight and comments on the project as it was ongoing. They would especially like to thank Kendra Marks for recruiting listeners for the study and coordinating the study.

Previous Presentations

Portions of this work were presented at the 177th Meeting of the Acoustical Society of America in Louisville, Kentucky. The meeting was in May, 2019.

References


Acknowledgments

The authors would like to thank the members of the Hearing Aid Laboratory for their insight and comments on the project as it was ongoing. They would especially like to thank Kendra Marks for recruiting listeners for the study and coordinating the study.

Previous Presentations

Portions of this work were presented at the 177th Meeting of the Acoustical Society of America in Louisville, Kentucky. The meeting was in May, 2019.
12 Seldess Z. MIAP: manifold-interface amplitude panning in Max/MSP and pure data. Paper presented at the Audio Engineering Society Convention 137; 2014
16 Wendt K. Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensität - un Laufzeitstereophonie [Directional hearing with two superimposed sound fields in intensity- and delay-different stereophony]. Aachen: Technische Hochschule; 1963