Acuity Immersion with e-STAT AI

Joyce Rosenthal and
Simon Carlile, Ph.D.

Introduction

Acuity Immersion takes the features of sound that enhance spatial hearing and makes them accessible to hearing aid wearers with mild-to-moderate hearing loss. Access to these features is what gives listeners their perception of sound source location, as well as their physical sense of presence, or immersion, in the environment. Complete spatial immersion is not possible without the shaping of sound by the pinna. With its complex contours, the pinna makes a unique acoustic mark on every sound that enters the ear. It is these acoustic signatures that not only help to disambiguate the location of a sound source, but also allow the listener to experience the uniqueness of each sound space.

What is spatial hearing?

Localization: Interaural Time and Level Differences

Open any textbook on auditory perception and there will be a chapter on spatial hearing. Even before Lord Rayleigh's research was published in Philosophical Magazine in 1907 (Rayleigh), the fact that we had two ears was recognized as important in helping us localize the source of a sound in the space around us. Two ears provide the auditory system with the opportunity to simultaneously sample a sound at two separate locations in the sound field.

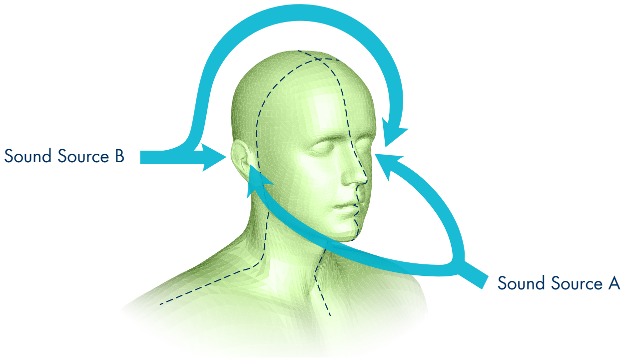

Figure 1. A sound coming from source A will have the same level, and same time of arrival at both ears (ITD=0, IDL=0). A sound coming from location B will be louder, and arrive sooner at the right ear. (ITD ≠ 0, IDL ≠ 0).

Since there is distance between the ears (the distance across the head), there will also be a difference in the time of arrival of sound to each ear. This is referred to as the interaural time difference (ITD). Of course, when the source is on the midline plane (Figure 1, sound source A) sound will arrive at both ears at the same time and the ITD will be zero, as the distance from the source to each ear is the same. When the source is located opposite one ear along the interaural axis (Figure 1, sound source B), the time delay to the other ear will be maximal and is determined by the diameter of the head. The average adult head provides for an ITD of around 600 to 700 μs, and amazingly, we are sensitive to changes in the ITD of the order of 10 μs (Zwislocki & Feldman, 1956).

A second difference in the sound received at both ears results from the head being relatively dense compared to the air. This causes sound to be reflected and refracted around the head, which changes the relative levels of the sound in each ear. These are referred to as the interaural level differences (ILDs). Naturally, when the sound source is on the midline plane, the paths from the source to each ear are the same and there is no ILD. For off-midline sound source locations, ILDs are small for low-frequency sounds and large for high-frequency sounds. This is because the wavelength of a low-frequency sound is much longer than the dimensions of the head, making the head an ineffective baffle. The head can more effectively block high-frequency sound with wavelengths much shorter than the dimensions of the head. Acoustic measurements show that the maximum ILD occurs for a location around 60° from the frontal midline and not, as one might think, for a location on the interaural axis (Shaw, 1974).

One of the reasons that the greatest ILD is not found on the interaural axis is that the outer ear, the pinna and the concha amplify the sound. The acoustic axis of that directional amplifier is directed to the frontal quadrant on the same side. At the same time, the complex physical structure of the pinna and the concha also filter the sound in a location-dependent manner. For any one location, this gives rise to level variations for different frequencies of more than 10 dB and the exact pattern of this filtering changes with the relative location of the source. These variations are referred to as the spectral, or monaural, cues to location and have been demonstrated to aid in monaural localization of the source (Slattery & Middlebrooks, 1994).

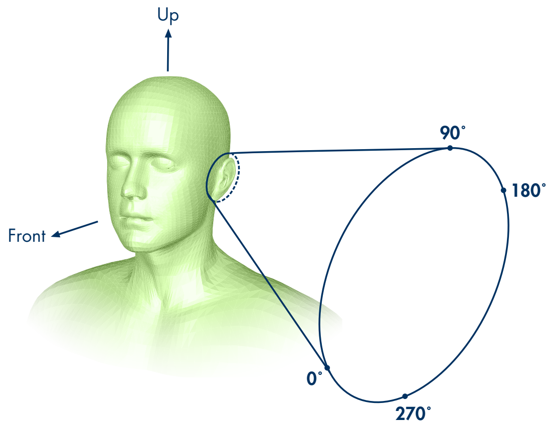

Figure 2. The "cone of confusion." ITDs and IDLs are ambiguous for sound sources located along the circumference of a cone centered along the interaul axis.

Localization: The Pinna Can Help When ILDs and ITDs Can't

Returning to our textbook account of spatial hearing, the ILD and ITD cues are often referred to as the binaural cues to sound location. Lord Rayleigh is usually credited with the first description of this "Duplex Theory" of localization (Rayleigh, 1907). The spectral cues of the pinna are incorporated into the Duplex Theory to help explain how the auditory system resolves spatial ambiguity in the binaural cues. Ambiguity arises because of the geometry of the placement of the two ears. This can be easily understood by considering the case of a sound located on the midline plane directly in front of the listener. As discussed above, this case will result in zero ILD and ITD cues, as the paths from the source to the two ears are identical. Consider, though, that the same situation arises for a sound directly above the head and directly behind the head. In fact, any location on the midline plane will give rise to zero ILD and ITD, making the binaural cues spatially ambiguous. This condition generalizes for each binaural interval so that, for instance, a sound located in front at 30° to the right of the midline will give rise to the same ITD as for a location 30° from the posterior midline, or indeed from any location described by the surface of a cone centered on the interaural axis. This is often referred to as the cone of confusion (Figure 2).

The cone of confusion is slightly more complicated for the ILD cue because of the direction of the acoustic axis of the outer ear, but generally the same situation arises. However, since the across-frequency filtering effects of the pinna vary with the location of the sound source, the pinna can provide the critical cues needed to resolve ambiguous ILDs. Figure 3 (modified from Carlile, 2014) shows an example of the pinna's filter effects measured at points along the midline plane (front to back). Frequency is shown on the x-axis, midline location on the y-axis, and color indicates the corresponding amplitude of the filter effect at any given frequency-location coordinate. Effects range in amplitude from +12 dB (dark red) to –15 dB (dark blue). As an example, the black arrow in

Figure 3 points to a blue region, indicating that at 8 kHz, when the sound source is located in front of the head, the pinna effect is to attenuate the sound by 15 dB. Note also that the biggest effects (bright reds and blues) occur over a range of 4 to 16 kHz. It is the combination of the precise binaural cues, together with the coarser spectral pinna cues, that provide for the accurate spatial localization abilities that we see in humans.

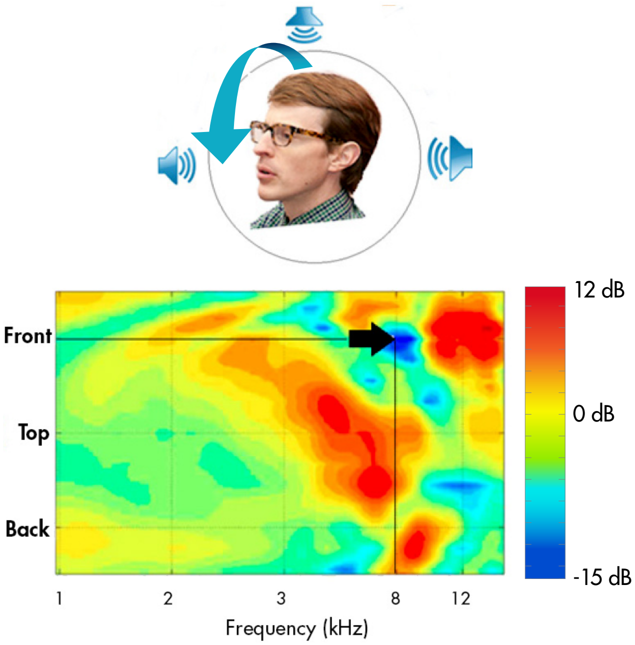

Figure 3. The variation in the pinna effects for one ear is shown as a function of location along the midline (from front to back). The colors of the contours indicate the amplitude of the function at a specific frequency (x)/ location (y) coordinate and extend from 12 dB (dark red) to –15 dB (dark blue). The black arrow points to a blue region indicating that at 8kHz the pinna attenuates sounds coming from the front midline.

Spatial Hearing: More than Just Localization

It is at this point that the textbook account of spatial hearing usually ends, leaving us with the impression that spatial hearing is all about localization. There is more recent research, however, that explains another aspect of spatial hearing. This research is mainly in the area of virtual reality, so it doesn't tend to show up in the traditional perceptual science textbooks. There is strong evidence that the spectral cues from the pinna are a very important component in generating the perception of sound outside the head and away from the body, which is referred to as a sense of externalization (Xie, 2013). This is demonstrated in the everyday experience of listening to music with earbud headphones. When sound enters the ear canal without being filtered by the pinna and the concha, the sound is perceived to be in and around the head. This is in stark contrast to the listening experience of music coming from speakers across the room. In this situation, the music is clearly perceived to be coming from a source away from the body. The difference between these two listening experiences is the filtering by the outer ear. The brain uses the outer ear-filtered signatures, or spectral cues, to generate the externalized percept.

Externalization also produces the sense of presence. Presence is often described as the feeling of being in the world. Of course, virtual reality research is focused on figuring out precisely what makes a simulated world feel real. This research has provided some important insights into what most of us take for granted: our sense of presence. It turns out that one of the most powerful drivers of the sense of presence in a virtual world is the rendering of 3D audio that the listener can interact with in a natural way. For instance, if one listens to a stationary sound source in a virtual environment while moving one's head, the source has to appear to move to maintain its position relative to the head. This doesn't normally happen when one listens to a sound over headphones since the directional reference is the ears. In this case, the sound source moves with the ears/head and is no longer perceived as being stationary in the external world.

Turning back to our example of listening to music from the speakers, the same music will be perceived differently depending on the room in which it's played. The listener will perceive the sound of the space as being small or large, reverberant or anechoic. The unique acoustics of each space give rise to our sense of spaciousness. This sense of spaciousness not only provides environmental context for sound, but also drives the emotional impact of the sounds we hear. Architects have known this for hundreds of years. The designs of some imposing spaces, like cathedrals and symphony halls, have taken advantage of specific acoustic effects to elicit emotional ones. The design of some listening spaces, like Carnegie Hall, do this much better than others.

Spatial Hearing and Hearing Loss

Based on what we've learned in previous sections about the pinna's contribution to spatial hearing, it's reasonable to predict that hearing loss could significantly impact a listener's spatial hearing experience. There is much research demonstrating that hearing impairment degrades localization ability (Akeroyd & Whitmer, 2011). High-frequency hearing loss will reduce the bandwidth of the outer ear's spectral cues. This, in turn, will decrease the brain's capacity to resolve otherwise ambiguous localization cues. We see this in the laboratory where, in the absence of any visual cues, hearing impaired listeners make many more front-back errors in locating a sound (Akeroyd & Whitmer, 2011; Best et al., 2010). It follows from the teachings of virtual reality research that reducing the fidelity and bandwidth of the spectral cues should also impact the sense of externalization. Although very little research has been done in this area, it has been shown that hearing impaired listeners rate the externalization and distance of sound sources very differently from normal hearing listeners (Boyd, Whitmer, Soraghan, & Akeroyd, 2012).

Acuity Immersion with e-STAT AI

So, what happens when information lost because of hearing impairment is reintroduced? Can the sense of presence and immersion be restored? The spectral cues for spatial hearing that are diminished with hearing loss can be made accessible again by 1) increasing bandwidth with amplification, 2) "warping" the cues to a different frequency region where there is better hearing or 3) as is done with Acuity Immersion, implementing a combination of both.

Starkey's Acuity Immersion with e-STAT AI is designed to improve the "sense of presence" for individuals with mild-to-moderate, high-frequency hearing loss. The feature is unique to CIC and IIC hearing aid styles. Essentially, Acuity Immersion with e-STAT AI takes spectral cues in the 4 to 10 kHz range, compresses them into the 4 to 7 kHz range and makes them audible. Unlike other frequency compression algorithms, which are designed to enhance speech cues, Acuity Immersion explicitly targets spectral cues and leaves speech sounds largely unaffected.

Over a period of a few weeks, the brain is able to use these newly accessible spectral cues to relearn the relationship between sound and space and ultimately restore the user's natural sense of presence. While research on the underlying mechanisms of spatial "remapping" and recalibration continues (Majdak, Walder, & Laback, 2013), there is clear behavioral evidence that relearning does occur (Mendonca, 2014).

A component in the success of the Acuity Immersion feature is that it's offered only on deep-fitting, custom hearing aids. Since the microphone of a Completely-In-Canal (CIC) or Invisible-In-Canal (IIC) sits within the ear canal, the patient's natural pinna cues remain intact for all signals entering the hearing aid. In contrast, the Behind-The-Ear (BTE)/Receiver-In-Canal (RIC) microphone sits behind the pinna so sound entering the hearing aid has completely bypassed the pinna's spectral shaping. As a result, spatial hearing algorithms implemented on BTEs and RICs face the impossible task of having to simulate spatial cues. At best, these algorithms can partially restore natural pinna cues by using directional microphones to mimic a natural directivity pattern of the outer ear (Boyd, Whitmer, Soraghan, & Akeroyd, 2012), one that will not exactly match the listener's own ears.

The goal of Acuity Immersion is to optimize spectral resolution and audibility to give the patient access to as many spatial cues as possible. It does this by first dividing the signal above 4 kHz into very narrow frequency bands. When the bands are compressed, a large number of the critical spectral peaks remain intact. Finally, when the signal enters the compressor and gain is applied, a compensatory boost is given to channels in which new spectral information has become available. More will be explained about this in the next section on e-STAT AI.

e-STAT AI

e-STAT AI and e-STAT are Starkey Hearing Technologies' proprietary formulas. e-STAT uses NAL-NL1 as its starting point, with modifications empirically derived on the basis of patient preference for gain (Scheller & Rosenthal, 2012).

e-STAT AI is the fitting formula recommended for use with Acuity Immersion. It is based on the NAL-NL2 fitting formula and modified to give softer pinna cues the gain boost needed to maintain prescribed output from 4 to 7 kHz. Like e-STAT, e-STAT AI includes the minimize vent interaction (MVI) algorithm, designed to prevent interactions between hearing aid output and direct sound through the vent, which can result in poor sound quality. E-STAT AI prescribes more gain than e-STAT above 4 kHz, which is the critical frequency region for pinna cues.

Figure 4 shows a comparison of e-STAT AI and NAL-NL2 targets. Note that the targets are identical in an output view (real ear SPL view, left panel), but that e-STAT AI provides more gain than NAL-NL2 above 4 kHz (insertion gain view, right panel). The gain is increased to compensate for the drop in compressor input level that occurs between 4 and 7 kHz when Acuity Immersion is enabled. The amount of the increase is sufficient to equalize e-STAT AI and NAL-NL2 output for a speech-shaped input. It is important to note that when Acuity Immersion is enabled, the extra gain applied by e-STAT AI does not increase the risk of feedback! The changed frequency content of the output from 4 to 7 kHz destroys its correlation to the input, thus making the hearing aid highly resistant to any feedback above 4 kHz.

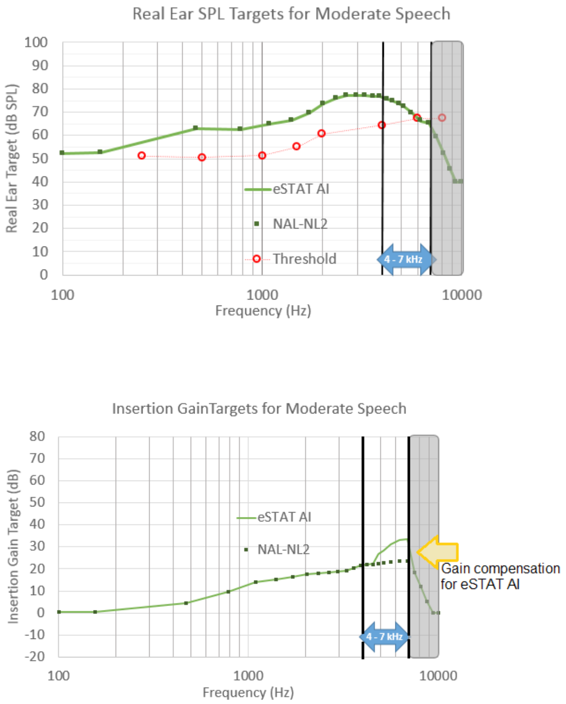

Figure 4. e-STAT AI and NAL-NL2 targets are the same in real ear SPL (left panel). To achieve the same output, more gain is needed for e-STAT AI. This is shown in the panel on the right. The additional gain is needed to compensate for a drop in input level to the compressor when Acuity Immersion is enabled.

Candidacy: Who Can benefit from Acuity Immersion with e-STAT AI?

While Starkey Hearing Technologies continues to recommend e-STAT as the starting point for a new fitting, an exception is made in the case of patients whose thresholds indicate the potential for benefit with the Acuity Immersion feature. These are patients with mild-to-moderate, high-frequency hearing loss who do well with amplified speech but may be experiencing a diminished sense of presence and difficulty separating sound sources in complex listening environments. For this targeted population, the additional gain provided by e-STAT AI was found not to change overall preference ratings for sound quality relative to e-STAT.

With the introduction of Acuity Immersion, there are now two features offered on CIC and IIC styles designed to make high-frequency spectral cues audible: Acuity Immersion and Speech Shift. Candidacy for each feature is based on audiogram: the main difference in criteria being the patient's threshold at 8 kHz. As a rule of thumb, patients with thresholds better than 65 dB HL at 8 kHz are considered candidates for Acuity Immersion. The Speech Shift candidate has a greater degree of hearing loss and more difficulty understanding speech. For the patient with this degree of hearing loss, the need for a normal sense of presence is secondary to the need for improved speech understanding. Audiometric criteria for these two features are mutually exclusive so that only one (or neither) will be enabled automatically by Inspire.

Summary

Acuity Immersion with e-STAT AI leverages the presence of natural pinna cues at the CIC or IIC microphone to improve sense of immersion for those individuals with mild-to-moderate, high-frequency hearing loss. It does this by

1. Increasing bandwidth for improved high frequency audibility,
2. Compressing high-frequency spectral cues in such a way that does not diminish speech quality, and
3. Boosting gain to compensate for what would otherwise be a drop in output with frequency compression.

Acuity Immersion is designed to reinforce the patient's sense of presence, the feeling of being in the world. Preservation of natural input, sufficient gain for audibility, minimal interference with speech and appropriate candidacy rules all work together to make Acuity Immersion with e-STAT AI the most natural sounding feature of its kind.

References

Akeroyd, M. A., & Whitmer, W. M. (2011). Spatial hearing and hearing aids. *ENT & Audiology News, 20, 76.*

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, 723–32.*

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 the Acoustical Society of America, 131, EL268-EL74.*

Carlile, S. (2014). The plastic ear and perceptual relearning in auditory spatial perception. *Frontiers in Neuroscience, 8, 237.*

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

Mendonca, C. (2014). A review on auditory space adaptations to altered head-related cues. *Frontiers in Neuroscience, 8, 219.*

Rayleigh, L. (1907). On our perception of sound direction. *Philosophical Magazine, 13, 214–32.*

Scheller, T., & Rosenthal, J. (2012). e-STAT fitting formula: The rationale behind the rationale. *Innovations, 2, 41–45.*

Shaw, E. A. G. (1974). The external ear. In W.D. Keidel & W.D. Neff (Eds.), *Handbook of Sensory Physiology: Auditory System.* (pp. 455–90). Berlin: Springer-Verlag.

Slattery, W. H., & Middlebrooks, J. C. (1994). Monaural sound localization: acute verses chronic unilateral impairment. *Hearing Research, 75, 38–46.*

Xie, B. (2013). *Head-related transfer function and virtual auditory display.* Plantation, FL: J. Ross Publishing.

Zwislocki, J., & Feldman, R. S. (1956). Just noticeable differences in dichotic phase. *Journal of the Acoustical Society of America, 28, 860–864.*

Starkey.com
@StarkeyHearing
facebook.com/starkeyhearing