Spatial Organization and Speech Understanding

Donald J. Schum, Ph.D.
Vice President, Audiology & Professional Relations
Oticon, Inc.

oticon
PEOPLE FIRST

SPATIAL ORGANIZATION and SPEECH UNDERSTANDING

Donald J. Schum, Ph.D.
Vice President, Audiology & Professional Relations
Oticon, Inc.

IMPROVING SPEECH UNDERSTANDING in noise continues to be the major focus of hearing aid development. If you ask an audiologist which factors account for a patient's ability to understand speech in a noisy setting, the response will likely include a discussion of factors such as audibility, signal-to-noise ratio, masking effects, frequency selectivity, etc. Although these factors are indeed relevant, they are focused on the signal conditions and the status of the peripheral auditory system. Further, "noise" is typically discussed as a unified concept: a totality of all of the sounds that are not the speech signal of interest.

If you ask the same question of a cognitive scientist, the answer will likely be quite different. The discussion will focus on how the listener deconstructs the totality of sound input into distinct sources, locates those separate sound sources in space and then focuses attention on the signal of interest while suppressing cognitive awareness of the competing sources. The discussion will focus on how the brain and binaural auditory system organizes the world of sound, allowing attention to be voluntarily directed at will. The term "noise" is not particularly relevant because the signal of interest versus the competition can vary from moment to moment, depending on the wishes of the listener. Multiple sound sources can be monitored at the same time, with the amount of cognitive capacity donated to one particular sound source under the control of the listener. A more appropriate description of the task would be speech understanding in complex environments.

Both descriptions are valid and relevant. However, hearing aid design has been, for years, driven by the audiological description of speech understanding in noise. We have developed compression, noise reduction and directional systems to best compensate for the disruptions in the peripheral auditory system, attempting to maximize audibility, control loudness and minimize the effects of noise. We have not necessarily viewed hearing aid development in relation to the cognitive process of listening in complex environments. We have tacitly assumed that if we optimized signal characteristics independently on each ear, we would maximize the ability of the auditory system to glean speech information.

What has not been part of the discussion over the years is how hearing aids can act to support the natural spatial organizational processes of the cognitive system. The brain thrives on broadband input and preserved spectral and timing differences between the ears to determine the location of sound sources. In our attempts to design hearing aids to optimize signal characteristics on each ear separately, we have, in some ways, disrupted some key information that underpins spatial resolution (Schum & Hansen, 2007). For the person with normal hearing, the brain can do an amazing job of finding and focusing in on the voice of interest in a complex listening environment in a way that no computer based listening system can. Consistent with the traditional medical philosophy of "first, do no harm", the primary goal of amplification should be to allow the listener to use as many preserved, natural abilities as possible. The role of the hearing aids should be to augment and supplement, making the best use of the remaining auditory capabilities of the user.

Hearing aid development has not been viewed in relation to the cognitive process of listening in complex environments.

THE USER RESPONSE TO EPOQ

We live in a complex world. We are constantly bombarded by stimulation. Humans have an amazing ability to sort through this onslaught and automatically and effortlessly make sense of the constant flow of events happening around us. Humans are extremely effective at identifying and monitoring target speech in a complex background of other talkers (Cherry, 1953). Epoq is a hearing system that was specifically designed to improve the ability of the patient to use the natural cues to spatial organization. The auditory system (two ears and the brain) constantly takes in a wide variety of cues and automatically, unconsciously and effortlessly recognizes not only what sounds are in the environment but also where those sounds come from. Epoq was designed to support that natural process.

Users are talking about aspects of their listening experience that we have not heard before.

Since its widespread introduction into the marketplace, we have been struck by the nature of what patients have been saying about their experience of hearing with Epoq. They are talking about aspects of their listening experience that we have not heard before. Here is an example from Asger, a 63-year old gentleman:

> *"I experience sound in a much broader, deeper and fuller way now. When I compare my old instruments to these new ones, it seems as if I have pressed the loudness-button on the TV. Suddenly everything has become more spatial and diverse. I can hear some extra dimensions that I couldn't hear before. These Epoqs have improved my hearing in every way – not just for music. This isn't an expression I would normally use, but they really do complete my hearing."*

Other users have used phrases such as "sound doesn't feel like it is right on top of me" and "I am hearing 360 again". As we have stepped back and taken a close look at what users like Asger are saying, we realize that Epoq has allowed patients to once again experience "Spaciousness."

THE BASIS OF SPATIAL ORGANIZATION

The human sensory systems function in such manner that we automatically and subconsciously use what ever information that we can gather to organize the world around us into meaningful objects and events. We normally are not consciously aware of the individual auditory, visual, olfactory or proprioceptive stimuli that are constantly bombarding us. Rather, we are aware of the objects and events all around us, especially those that are meaningful to us.

All sounds that we perceive have certain fundamental properties. Hearing care professionals are most used to talking about pitch and loudness. One fundamental property of sounds that we often overlook is location. All sounds come from ***somewhere***. Sound without location is not reality. Although we are usually not consciously aware, when a normally hearing person hears something they also have information about the location of the source of that sound (Blauert, 1997). Sometimes location is relevant, sometimes it is not. However, information about location is available if we need it.

Sound without location is not reality.

More importantly, when we are listening in challenging environments with multiple sources of sound, location is ***THE*** cue that we employ to arrange all of the sound sources within our environment and then focus on what we find most relevant. Scene Analysis refers to the cognitive activity of deconstructing the totality of sound input into isolated sounds sources arising from specific places in space (Bregman, 1990). It is the way we perceptually organize the world of sound around us into meaningful objects and events, with specific spatial relationships. We usually talk about paying attention to the voice of interest and ignoring everything else. The reality is that before we can ignore something, at some lower level of consciousness, we have to know what it is and, usually, where it is coming from. When a new sound occurs, it draws our attention until we can identify it and decide if is relevant to pay attention to. If it is relevant, we consciously monitor it. If it is determined to be irrelevant, it fades from our immediate consciousness.

CUES TO LOCALIZATION

Localization of sound sources on the horizontal plane is possible due to timing and level differences of sounds arriving at one ear versus the other (Yost, 1994). It has long been recognized that, when sounds originate somewhere other than precisely in front or behind us, the sound will reach the near ear earlier and at a higher level than the far ear. We use timing difference below 1500 Hz. and level differences above 1500 Hz. to determine where the sound came from. The timing differences are due to the finite speed of sound (sounds from 90 degrees to the side arrive at the near ear about .6 to .7 msec before arriving at the far ear) and level differences are due to the ability of the head to attenuate sounds with shorter wavelengths (by up to 20 dB). We are also sensitive to location of sounds that occur from somewhere other than the same elevation of the ears, using acoustical information in the extreme high frequencies (above 4000 Hz.) This skill is primarily related to the convoluted shape of the pinna, where differences in source elevation will interact with the wrinkles and folds of the external ear in such a way to create significant peaks and valleys in the extreme high-frequency response. Locating in the vertical plane is primarily driven by the comparison between direct sound in the ear canal and pinna reflections. Vertical localization is also affected by head and torso reflections by a comparison of direct sound to the ear canal versus sound generated from the head & torso reflections.

In terms of speech understanding improvement, the signal-to-noise ratio (S/N) can be improved by up to 5 dB if the noise is something other than a competing talker. If the competition is one or two other voices, the improvement can be on the order of 5 to 10 db or greater with 90 degrees of spatial separation (Arbogast, Mason & Kidd, 2005; Hawley, Litovsky & Culling, 2004; Kidd, Mason & Gallun, 2005). The extra improvement clearly has a cognitive component, showing the power of identifying different speech streams in space and focusing attention on one versus another.

REVERBERATION

One major component of our sense of spaciousness is reverberation. Reverberation refers to the sounds that we hear that do not travel directly to our ears but rather bounce off of one or more surfaces. It is extremely rare for us to hear only direct sound. We nearly always hear a combination of direct and reflected sound. In fact, sound without reverberation is usually perceived as being weak and thin. Reflections that arrive within the first 100 ms are usually incorporated into our perception of the direct sound, adding a sense of fullness to the sound. Sounds arriving later are perceived separately and heard as reverberation. Reverberation becomes a negative characteristic of a listening space when the later arriving reflected sound over powers the direct sound and early arriving reflections (Blesser & Salter, 2007).

Developing an aural representation of the space you are in is an essential component of being able to place all of the sound sources accurately within that environment.

An important facet of spatial hearing abilities is our perception of the characteristics of the space we are in. If you walk blindfolded into an unknown room, it takes only a few moments of listening to sound in that space to develop an impression of where you are. If indoors and someone is talking, you immediately have a sense for the dimensions of the room. You can tell something about whether or not the space has carpet or hard-surfaced flooring. You can tell how far the talker is from you. Developing an aural representation of the space you are in is an essential component of being able to place all of the sound sources accurately within that environment.

When evaluating how people describe the sound of spaces, professionals in the field of Architectural Acoustics have closely examined the words people use (Blesser & Slater, 2007). One of the fundamental terms that has emerged is the sense of "spaciousness". This descriptor captures the breadth and depth of the setting where listening takes place. In a similar vein, the Swedish researcher Gabrielsson and his colleagues spent a number of years investigating the words people use to describe sound (e.g., Lundberg et al., 1992; Ovegard et al., 1997). They were interested in developing scaling materials for sound reproductions systems, including loudspeakers, headphones and even hearing aids. They also identified "spaciousness" as one of the core dimensions of sound perception.

It turns out that our sense of "spaciousness" is determined by our ability to localize not just direct sound but also reflections. Normal hearing listeners are sensitive to both the timing and angle of arrival of these early reflections. We can tell a lot about the space we are in based on when these reflections get to us (compared to the direct sound) and the direction from which they arrive. The better our localization abilities, the greater our sense of spaciousness.

LISTENING IN ALL DIMENSIONS

Although we tend to study the auditory system by looking at one discreet, well-controlled dimension or task at a time, perception in normal life is far more complex. The brain will grab hold of any information that is available to organize all of the sounds around us. The brain will take in information on horizontal location, vertical location, movement, reverberation, setting, visual cues, expectations, familiarity with the talker, etc. in order to decipher the speech of interest.

The brain will grab hold of any information that is available to organize all of the sounds around us.

No one dimension of spatial organization is predominant or essential for accurate perception. The robustness of any given cue will vary depending on many factors. Perhaps I am looking down at the newspaper while having a conversation and the talker is both to my side and higher than the level of my ears. Perhaps the setting is too reverberant for that to be used as a cue. Perhaps the talker is too far away for visual cues to be helpful. For amplification to support natural spatial organizational skills, it must faithfully preserve as much natural information in the environment as possible.

WHAT DOES EPOQ DO TO SUPPORT SPATIAL RESOLUTION?

The best way to answer this question is to first ask what the brain needs in order to organize the auditory world. The brain relies on information that is supplied by the two separate ears. Sometimes the signal of interest is in one ear and not the other, for example while listening on the phone. However, it is more common that auditory signals are present in both ears, but not necessarily in exactly the same form. The brain also looks for relationships between the signals at the two ears (Yost, 1994). Synchronized differences between the two ears provide crucial information as to where the sounds come from. Being able to assign location to all of the sounds in the environment is the first step in separating the signal of interest (especially speech) from background competition. Preservation of ear-to-ear differences is essential in amplification in order to provide the brain with all of the information that it needs. In addition, accurate preservation of information in the higher frequencies is essential for localization in all dimensions.

Epoq is a new advanced technology hearing aid by Oticon. Along with a variety of other signal processing improvements, there are specific features in Epoq designed to improve spatial resolution for the user (Schum, 2007).

» *Wireless Communication Between Hearings Aids:*

A version of digital magnetic wireless communication is implement in Epoq at the chip level to allow for high-speed data transfer between the two hearing aids in a binaural fitting. This feature allows for the coordination of switch position and program selection, synchronization of the action of noise reduction and adaptive directionality, improved accuracy of feedback and own voice detection and, most importantly, preservation of naturally occurring interaural spectral differences. As indicated above, these spectral differences in the mid and high frequencies form the basis of horizontal localization. The high-speed sharing of data between the hearing aids allows us to implement a feature called Spatial Sound, which adjusts the gain and compression response of the hearing aids to preserve these important cues.

» *Extreme Bandwidth:*

The bandwidth of Epoq is an industry leading 10 kHz. In all models. Pushing the response of the device out to this extreme allows for the preservation of localization cues that have historically been inaudibility in previous generations of hearing aids.

» *RITE Design:*

The movement away from traditional earmolds and towards thin tube and receiver-in-the-ear (RITE) fittings in combination with ever-improving feedback cancellation algorithms has put new emphasis on the advantages of open fittings. Beyond the minimization of occlusions, open fittings also allow for the passage of high-frequency information even beyond the bandwidth of the hearing aid. Again, the better the perception of high-frequency information, the greater likelihood of effective spatial resolution. Further, compared to thin tube approaches, placement of the receiver in the canal provides a smoother response, improving sound quality, decreasing feedback risk and maximizing the ability to provide complete high-frequency gain.

Figure 1 demonstrates the effect of preserving interaural intensity differences in the Spatial Sound feature of Epoq. Recordings were made in the ear canals of KEMAR with sound presented 90 degrees to the right of midline. The stimulus was white noise presented at 75 dB SPL. Recordings were made under three conditions: (1) open ear canal, (2) binaural Epoq fitting with Spatial Sound feature disabled and (3) binaural Epoq fitting with the Spatial Sound feature enabled. For these recordings, the directional and noise reduction systems were disabled. The Epoq RITE hearing aids were fit for a moderate, high frequency hearing loss and an open dome was used. All of the recordings were then bandpass filtered from 1 to 10 kHz. (to compare the response within the functional bandwidth of the device).

In Figure 1, the left-most panel shows the recordings in the near and then far ear canal. The overall difference from one side of the head to the other was 8 dB, reflecting the average interaural spectral difference at + 90 degrees. The middle panel shows the effect of compression when operating independently on one ear versus the other. The natural 8 dB interaural difference has now been reduced to 2.5 dB. The right-most panel shows the effect with Spatial Sound activated. The gain and compression response has been adjusted to retain interaural spectral differences. In this case, the difference from the right to the left ear is 6.5 dB. When Spatial Sound is active, the input level to both ears is shared between the hearing aids and the gain/compression response is based on a central decision necessary to, as best as possible, maintain naturally occurring interaural spectral differences.

Taken together, the features of Spatial Sound, Extreme Bandwidth and RITE design have placed emphasis of feeding the brain the most complete, best preserved information that forms the basis of spatial resolution. As discussed above, in attempts to maximize audibility and minimize the effects of background noise, hearing aids have been limiting or distorting the natural cues to localization.

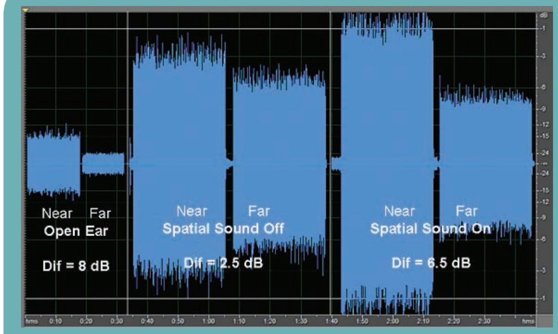

Figure 1: *Recordings of white noise made in the ear canals of KEMAR. The speaker was oriented at +90 degrees to the right of midline and recordings were made in both the near and far ears. Recordings were made with the ear canals open (left panel), with Epoq RITEs in place with the Spatial Sound feature disabled (middle panel) and then enabled (right panel).*

MEASURING SPATIAL PERCEPTION

We do not measure localization in routine clinical practice. Accurate objective measurement of localization requires speaker arrangements and support electronics that simply are not available in typical hearing aid facilities. Further, measuring the ability to localize discrete signals in sterile test environments likely does not capture the user's experience of sound in space (especially in everyday environments).

Traditional measurements do not capture the listener's experience of perceiving sound in space.

In order to assess a user's perception of sound within the environment, Gatehouse and Noble (2004) developed the Speech, Spatial and Qualities of Hearing (SSQ) questionnaire. This scale is comprised of 53 items across three domains: Speech understanding (mainly in challenging, real-life environments), Spatial Perception (assessing perception of localization, distance, movement, etc.) and Sound Qualities (assessing the ability to separate sounds from each other, familiarity of known voices and other sounds, distinctiveness of common sounds, etc.). The SSQ is designed to assess those aspects of the listening experience that are driven by factors beyond just audibility and S/N. They have used the SSQ, for example, to demonstrate the advantages of bilateral over unilateral hearing aid fittings (Noble & Gatehouse, 2006).

EPOQ AND THE SSQ

In an attempt to verify the effectiveness of the specific technical features in Epoq designed to improve spatial perception, selected items from the SSQ were used as part of a larger clinical study of Epoq fittings (Hansen, in preparation). The SSQ results from 76 patients from that study are reported here. We selected 17 items from the SSQ that specifically focused on spatial resolution. Examples of some of the SSQ items used in the present investigation are shown in Table 1.

SUBSCALE	ITEM
Speech	You are with a group and the conversation switches from one person to another. Can you easily follow the conversation without missing the start of what each new speaker is saying?
Spatial	Do you have the impression of sounds being exactly where you would expect them to be?
	You are outdoors in an unfamiliar place. You hear someone using a lawn mower. You can't see where they are. Can you tell right away where the sound is coming from?
Qualities	Can you easily ignore other sounds when trying to listen to something?

***Table 1:** Example items from the SSQ used in the present study.*

Each patient was an experienced bilateral user of advanced technology, multi-channel nonlinear digital hearing aids, mainly Oticon Syncro (83%). As part of the current study, each patient was fit with one of five styles of Epoq: ITC = 9 , ITE = 20, BTE = 21, RITE with a custom earmold = 10, RITE with a open dome = 16. All patients were within the published fitting range of the particular model of Epoq used. In general, patients were fit with the same model of Epoq as their Syncro fitting. Since Syncro is not available in a RITE model, the 31 Epoq RITE users were drawn from users of other Syncro models. All testing with Syncro occurred before the patients were fit with Epoq. Patients wore the Epoqs for at least one week before testing began with that product.

The SSQ results for the Syncro and Epoq fittings are shown in Figure 2 as "Two" and "Epoq", respectively. As can be seen, there are consistent improvements when moving from Syncro to Epoq on these spatial dimensions, especially when the situations are more difficult than just listening to one person in quiet.

To provide some perspective on these SSQ improvements, Figure 2 also includes a line named "One". Nobel and Gatehouse (2006) provided normative data for the SSQ when comparing unilateral to bilateral fittings for a large group of patients. The difference between a bilateral and unilateral fitting from that normative dataset was used to construct the expected performance of patients if they were fit with one Syncro. That data is presented as "One" in Figure 2. As you can see, if you compare the benefit of a bilateral fitting over a unilateral fitting, as reported by Nobel & Gatehouse (2006), it is approximately the same magnitude the difference between a binaural Epoq fitting and a bilateral Syncro fitting. Essentially all clinicians will agree that a bilateral fitting is superior to a unilateral fitting on many important dimensions. Based on these data, going from Syncro, a top-of-the-line multi-channel nonlinear product, to Epoq provides a similar magnitude of spatial benefit as going from one to two hearing aids.

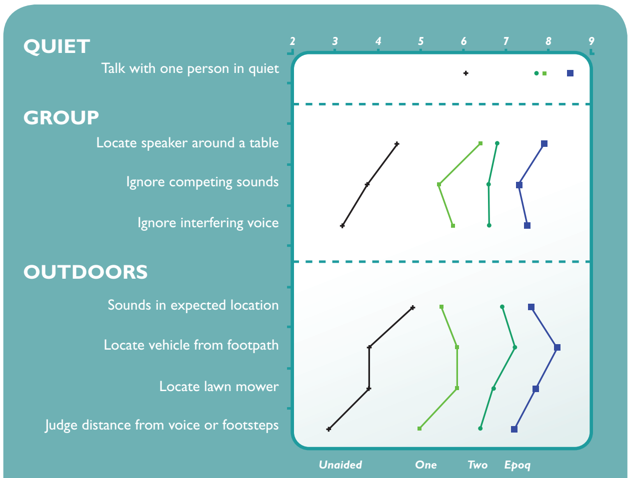

***Figure 2:** The group results on the SSQ for patients fit with Syncro ("Two") and Epoq ("Epoq") from the current study and for patients fit unilaterally ("One") from Nobel & Gatehouse (2006), adjusted to current technology (see the text for a further explanation).*

CONCLUSIONS

What does this improved sense of spatial perception mean to the patient? This renewed sense of spaciousness underpins the user's ability to organize the world of sound. The wall of sound that patient have been experiencing – the sensation that all of the sounds of the environment are mixed together and right there in front of them without any spatial dimension – is exploded. Sounds are dispersed back into the surroundings in their natural arrangement. With sounds back where they belong, listeners can now focus their attention on the signals of interest, as they now stand out from the cacophony of everyday life. Spatial hearing abilities provide the distinctness of individual sounds that is essential for effective hearing.

We see this effect when looking at the responses of the SSQ questionnaire. More importantly, we hear the effect when we listen to what users are saying:

"I hear background, but it is not bothersome"
"unwanted, extra sounds are lowered"
"noise is there, just less annoying"

And:

"I feel more comfortable in conversations."
"I can count on my hearing whenever and wherever I want to be."
"I can add to conversations without fear of making a fool of myself."

Our ability to hear, especially in a busy modern world, is a powerful skill. Epoq is specifically designed to allow the user to use his or her hearing as it is intended.

REFERENCES

Arbogast, A., Mason, C. & Kidd, G. (2005). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America. 117(4). 2169-2180.

Bregman, A. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press.

Blauert, J. (1997). Spatial Hearing. Cambridge, MA: MIT Press.

Blesser, B & Salter, L. (2007). Spaces speak, are you listening? Cambridge, MA: MIT Press.

Cherry, E. (1953). •Some experiments on the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America, 25, 975-979.

Gatehouse, S. & Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). International Journal of Audiology. 43, 85-99.

Hansen, L.B. (2007). The Epoq Product Test. Oticon A/S, in preparation.

Kidd, G., Mason, C. & Gallun, F. (2005). Combining energetic and informational masking for speech identification. Journal of the Acoustical Society of America, 118, 982-992.

Lundberg, G., Ovegard, A., Hagerman, B. Gabriellsson, A. & Brandstrom, U. (1992). Perceived sound quality in a hearing aid with vented and closed earmould equalized in frequency response. Scandinavian Audiology, 21, 87-92.

Noble, W. & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fittings on abilities measured by the SSQ. International Journal of Audiology, 45, 172-181.

Ovegard, A., Lundberg, G., Hagerman, B. Gabriellsson, A., Bengtsson, M. & Brandstrom, U. (1997). Sound quality judgment during acclimatization of hearing aid. Scandinavian Audiology, 26:43-51.

Schum, D. J. (2007). Redefining the Hearing Aid as the User's Interface with the World. The Hearing Journal, May.

Schum, D. & Hansen, L.B. (2007). New technology and spatial resolution, Audiology Online. www.audiologyonline.com/articles/pf_article_detail.asp?article_id=1854.

Yost, W. (1994). Fundamentals of Hearing: An Introduction. San Diego: Academic Press.

NOTES