Binaural Processing for Understanding Speech in Background Noise

How it works, why it doesn’t always work, and how we can make it better with binaural beamforming algorithms

Joel Beilin, MScEE, Rebecca Herbig, Au.D.

© Siemens Audiology, 2014.
Abstract

This paper describes the natural binaural hearing processes available to normal hearing individuals to improve speech understanding in noise, and why these processes are compromised in individuals with hearing loss. We also discuss binax beamforming solutions which help to compensate for these consequences and enhance binaural processing.

Methods to optimize speech-in-noise understanding

Since the introduction of digital hearing instruments, there are three primary methods to improve speech understanding in background noise: providing effective audibility, directional microphone technology, and digital noise reduction.

Introduction

We know that the majority of hearing impaired individuals seek amplification because of difficulty understanding speech in background noise. Improvement for this communication situation becomes a reasonable expectation following the purchase of hearing instruments, and satisfaction with amplification usually is then driven by the degree that this expectation was met. Not surprisingly, therefore, a long-standing challenge in the design, manufacturing and fitting of hearing instruments is improving the understanding of speech in background noise.

So how are we doing? Unfortunately, surveys suggest not great. The most recent MarkeTrak survey (Kochkin, 2010) revealed that while a respectable 91% of hearing instrument users are satisfied with their hearing instruments' performance for one-to-one conversations, only 37% were either “satisfied” (26%) or very satisfied (11%) with the performance of their hearing instruments for understanding speech in background noise. Progress in this area appears to be slow, as this 2010 satisfaction rate is only 8% higher than the 2004 MarkeTrak data (Kochkin, 2004).

Failure to meet the users’ expectations usually results in reduced hearing instrument use. About 8-10% of hearing instrument owners in the U.S. never wear their hearing instruments. The leading reason for this, given by 55% of respondents, is poor overall benefit, and benefit in background noise. This accounts for around 600,000 hearing instrument owners, who are non-users in the U.S alone.

It is certainly possible, that the expectations of many new hearing instrument wearers are unrealistic regarding the speech-in-noise benefit provided by hearing instruments, which then in turn leads to low satisfaction. But, nevertheless, it remains critical to optimize performance for these listening situations using whatever fitting techniques and technology that is available.

Effective audibility

There are two basic approaches for improving speech understanding in background noise. One is to attempt to reduce or eliminate the noise. A second approach is to emphasize the speech; that is, maximizing the amount of audibility that contributes to speech understanding. While we normally use both approaches simultaneously, the importance of effective audibility is sometimes overlooked. An example of this was shown in the research of Leavitt and Flexer (2012). They examined the speech-in-noise performance of experienced hearing instrument users using premier hearing instruments programmed to the manufacturer's proprietary algorithm versus the same hearing instruments programmed to the NAL-NL1 prescriptive targets. The QuickSIN SNR average score for all manufacturers was at least 4-5 dB better for the NAL-NL1 gain algorithm, and for two manufacturers, the average improvement was 9-10 dB SNR for the NAL-NL1. This degree of improvement in SNR is greater than what could be achieved with any noise reduction or directional technology alone, and emphasizes the importance of effective audibility.

There is an important difference, however, between optimizing audibility and maximizing audibility. Too much audibility in certain frequency regions can reduce sound quality, listening comfort, and cause hearing instrument rejection. A compromise is necessary. Siemens has found this compromise through the use of the proprietary fitting algorithm, binax fit. Independent research has shown that the binax fitting algorithm, which is based on micon fit, provides good “first acceptance” for the patient, and has been rated highly when compared to the proprietary algorithm of other manufacturers (Powers and Beilin, 2013). Most importantly, research at Vanderbilt University has found that for speech understanding in background noise, this fitting algorithm is equal to the widely-accepted NAL-NL2 prescriptive method (Powers et al, 2014).

Directional technology

Siemens has been a leader in directional technology since it was introduced in the early 1970s. Historically,
this technology has been the best solution for improving the SNR using conventional wearable hearing instruments. Over the past 15 years, clinical studies using Siemens hearing instruments have supported the patient benefit that can be obtained from dual-microphone directional instruments (Ricketts and Dhar, 1999), the three-microphone directional system (Bentler et al, 2004), adaptive directional microphones (Ricketts et al, 2005), and an anti-cardioid directional algorithm (Mueller et al, 2011). In general, this research has shown a 4-5 dB SNR advantage for directional technology in a diffuse noise field. This advantage is even greater for the Siemens adaptive polar pattern when single-source noises are present. Ricketts et al (2005) report a substantial advantage (compared to omnidirectional) of 8.5 dB for a stationary noise, and 7.5 dB for a moving noise source when the adaptive algorithm was employed.

Digital noise reduction (DNR)

The primary purpose of applying digital noise reduction (DNR) is to reduce the annoyance of background noise. This provides more relaxed listening, enhances ease of listening, and therefore, has the potential to indirectly improve speech understanding. The listeners are more apt to use their hearing instruments in difficult listening situations, and if they are more relaxed, it is more probable they will “tune in” to the conversation. Research with DNR also has shown that this technology may reduce the cognitive load, which allows for enhanced dual processing and in particular for improving short term memory associated with communication. One of the only studies showing a direct improvement in speech understanding with DNR was published by DiGiovanni et al, (2012) using Siemens technology. They found a 12% improvement in speech understanding with DNR “on” versus “off.”

Combining directional and DNR

One reason why substantial improvement has not been observed for speech understanding in background noise when using DNR is that there must be a careful balance between reducing noise, but not reducing important speech cues. As the DNR becomes more aggressive, the speech signal easily can be distorted. Of course, if the hearing instrument could identify portions of the incoming signal that did not contain the target speech, such as signals from behind the hearing instrument user, then DNR could be made more aggressive for these inputs. This is exactly what Siemens employed with Directional Speech Enhancement (DSE), an algorithm that combines DNR processing with directional technology.

The Siemens DSE uses directional technology to identify the location of the target speech signal, and then increases the strength of the DNR for other spatial areas where speech is not of interest. For example, when speech is from the front, the DNR will be more aggressive for multitalker babble from the back than from the front, resulting in an improved SNR at the ear. This benefit was documented with research at the University of Iowa, which showed a 6.5 dB SNR advantage (compared to omnidirectional) in a diffuse noise listening situation (Powers and Beilin, 2013). An SNR improvement of this magnitude could result in improved speech understanding of 50% or more. As reviewed here, technology has advanced significantly over the years, and we are now able to provide substantial SNR benefit for many listening situations. However, going back to the MarkeTrak data, it appears that even more technological advances are needed to meet the listening requirements of the hearing impaired. It is reasonable, then, to go back to the factors that contribute to understanding speech in background noise for individuals with normal hearing. One of these is binaural processing.

Binaural processing and SNR improvement

One method that can be used to improve speech understanding in background noise is not contained within a hearing instrument—it’s something the patients bring with them to the fitting—their binaural processing abilities. We know that binaural hearing is critical for auditory localization, for the spatial perception of the auditory scene, and is helpful for sound quality and general environmental balance. Binaural hearing, however, also improves the signal-to-noise ratio when listening to speech in background noise. We’ll review three factors related to this for individuals with normal hearing.

Binaural redundancy

Stated simply, the brain has two chances to get it right (Figure 1). The signal from the right and left ears are processed, and if two very similar signals are available (i.e., the desired speech signal), there is a greater probability that a correct identification will be made. Moreover, when only partial information is available, the redundancy allows for utilization of the best information from either ear. This can provide a 1-2 dB improvement in the SNR.
Binaural squelch

Through the use of phase, spectral and intensity differences, the auditory centers of the brainstem compare the signals from the two ears, and are able to give greater emphasis to the meaningful signal (Figure 2). This is commonly referred to as binaural “squelch,” as the unwanted signal (e.g., noise) is not given the same degree of neural processing, and is therefore somewhat diminished. This selective processing can provide a 1-2 dB improvement in the SNR.

Directed binaural listening

When signals from both ears arrive at the auditory cortex, two important functions that occur are fusion and separation, also known as segregation. By fusion, we mean that the brain can analyze each individual signal, and then fuse the two signals together in a manner that results in an outcome better than either signal alone; related to the redundancy effect discussed earlier. By separation or segregation, we mean that the brain can analyze the combined signal from the two ears, and pull out individual acoustical ‘objects’ from both signals, or in some cases, only the signal of interest. This later cortical function is important for understanding speech in background noise. Through a conscious effort, we can focus on a signal from one ear or the other, or spatially separated signals, while suppressing the signals of non-interest. This SNR advantage has not been precisely defined, but it is much larger than the two previously mentioned factors, which occur without conscious direction.

Effects of hearing loss and aging on binaural processing

We have described three binaural factors that can function to improve the SNR, and enhance speech understanding. These factors have been studied extensively in younger listeners with normal hearing. But how does bilateral sensorineural hearing loss impact this binaural processing? Or, considering that most hearing instrument candidates are over the age of 65, how do the cognitive deficits associated with aging interact with binaural processing?

Binaural redundancy

In a typical downward sloping hearing loss, much of the speech signal for average speech is no longer audible. A person who once had a speech intelligibility index (SII) of 1.0 may now have an SII of 0.50 or poorer. Simply stated, the redundancy is not as redundant, as large portions of the signal are not reaching the central auditory processing centers. Moreover, cochlear distortions often reduce the fidelity of the portions of the signal that are audible.

Treatment with traditional amplification: Certainly, well-fitted hearing instruments will improve audibility, and the degree of redundancy will be greater than without amplification. However, for most hearing losses, a significant portion of the speech signal remains inaudible, even with bilateral hearing instruments. Consider the patient whose hearing loss drops to 60-80 dB for 3000 Hz and above—a common audiogram for a new hearing instrument user. Even when fitted to a validated prescription such as the NAL-NL2, this patient will still only have an SII of 0.70 to 0.75, considerably less than someone with normal hearing. Moreover, amplification cannot correct the cochlear distortions that may be present.
Binaural squelch

As with the redundancy effect, the ability of the brainstem to squelch certain signals is somewhat dependent on the audibility of these signals. Reduced audibility is reduced squelch. Additionally, the phase and intensity comparative processing that works effectively when hearing is normal, typically is not as effective when the signals are asymmetric; rarely do individuals have the exact hearing loss and loudness growth in each ear. Cochlear distortions associated with sensorineural losses also can negatively impact binaural squelch.

Treatment with traditional amplification: The increased audibility from hearing instruments may assist the squelch effect, but it is unlikely that it will be as effective as with normal hearing individuals. Portions of the speech signal will still be inaudible, and it’s probable that the aided signals will be somewhat asymmetrical, impacting the intensity and phase comparisons. And in fact, some aspects of signal processing may alter phase relationships.

Directed binaural listening

While audibility and fidelity also are important for directed listening, this processing is commonly affected in older individuals by cognitive and central auditory processing deficits. As the central auditory system ages, cortical separation ability is diminished. Moreover, even mild cognitive disorders can compromise the patient’s ability to focus on the desired signal, when competing similar signals are present.

Treatment with traditional amplification: It is of course a prerequisite for cortical auditory separation that the target speech signal is audible. However, when the competing speech signal is provided the same amount of gain boost, the overall increased audibility is not very helpful for someone with an auditory processing or cognitive deficit.

Binaural beamforming: The binax solution

It’s clear, that we could improve speech understanding in background noise for the hearing impaired if we could restore the binaural processing capabilities that are present for young individuals with normal hearing. Unfortunately, we can’t—cochlear hair cell damage, central auditory neural dysfunction and cognitive deficits are not something that can be restored. Rather, what we can do is develop alternative means to compensate for these binaural auditory shortcomings. This is the starting premise for the development of the Siemens binax binaural beamforming processing scheme.

Through the wireless transmission and sharing of audio information between the two hearing instruments with e2e wireless 3.0, it is possible to make calculations concerning the auditory scene that were not previously possible. In particular, to focus and enhance the hearing instrument output for signals of interest—most commonly target speech—while simultaneously minimizing the output from other spatial orientations where the input is not desirable. This is accomplished through two different algorithms referred to as Narrow Directionality and Spatial SpeechFocus.

With Narrow Directionality, the directional polar pattern has a very narrow focus to the front (e.g., the look-direction of the hearing instrument wearer). The output of the hearing instruments is significantly reduced for all inputs falling outside of the narrow focus (Figure 4), even if they are located right next to the target speech source. This provides an SNR advantage not observed in previous directional instruments, and easily achieved simply by the listener looking at the talker of interest.

![Figure 4. Compared with standard monaural directional microphone (gray-shaded area), Narrow Directionality (purple-shaded area) has a narrower focus beam so that sounds outside of what is immediately in front of the wearer can be attenuated.](image)

On the other hand, the Spatial SpeechFocus algorithm is designed for the opposite speech-in-noise listening situation—when the listener cannot look directly at the target speech signal. This algorithm detects the location of speech (front, back, or either side) and adjusts the focus of the directional polar pattern to enhance this signal, which providing less output for signals from other spatial orientations (Figure 5).
feature, which automatically directs the focus of the directional polar plot to achieve maximum output only for the target speech signal, while signals from other spatial orientations receive less gain, therefore improving the overall SNR.

Summary

In summary, while we are aware that we cannot restore optimal binaural processing for older individuals with hearing loss, there are methods to compensate for their diminished capabilities. The new algorithms of the binax hearing instrument, using binaural beamforming, can provide this compensation very effectively; to the extent that the hearing impaired can perform even better than their normal hearing counterparts for some speech-in-noise listening conditions.

References


Kochkin, S. 2010 MarkeTrak VIII: Customer satisfaction with hearing instruments is slowly increasing. The Hearing Journal, 63 (1),11-19.


1 Wilson et al (2012), compared individuals with hearing loss to those with normal hearing, and found on average: For the BKB-SIN, the hearing-impaired group needed a 4 dB SNR advantage to perform equal to the normal-hearing group. For the QuickSIN, which uses more difficult speech material, an average 8 dB SNR advantage was required. Because many patients with hearing loss expect performance equivalent to normal hearing when wearing hearing instruments, these data remind us that we must not only restore audibility, but also improve aided signal-to-noise ratio (SNR) considerably if this level of speech understanding is to happen—a daunting task.

2 Many speech-in-noise tests are geared to find the point where the patient/subject can recognize (understand) 50% of the words. A signal-to-noise ratio is then recorded for that point. This is often referred to as the SRT-50 or SNR-50. Two common tests used for this purpose are the Quick Speech-In-Noise (QuickSIN) and the Hearing In Noise Test (HINT). The scoring of the QuickSIN takes the performance of normal hearing individuals into account, and the score is then reported as SNR Loss, rather than SNR-50. In research originating from Germany, the commonly used SNR-50 test is the Oldenburger Satztest (OLSA). When using any of these SNR-50 tests to define performance, negative numbers are good. For example, an SRT-50 of +2 dB would mean that the speech signal had to be 2 dB louder than the noise for recognition, whereas an SRT-50 of -2 dB would mean that the person could recognize 50% of speech when the noise was 2 dB louder than the target speech signal.

3 While many speech-in-noise tests are geared to obtain the SNR-50 point for the patient, and the score is reported as such, this value is not meaningful to most. For example, the average patient doesn’t really know if a 4 dB SNR improvement is good, bad, or just average. It is common, therefore, to convert dB SNR values to percent improvement, something that most everyone understands. This is only an approximation, however, as the importance of a 1 dB SNR change varies as a function of the steepness of the patient’s SNR performance/intensity function; a reflection of the difficulty of the speech material and the energetic/informational masking of the background noise. An average value, however, of 10% dB commonly is used depending on the particular speech test applied. Importantly, this improvement also only applies when the listening situation is around the patient’s SNR-50 point. For example, we might think of hearing instruments providing a 6 dB SNR benefit as giving the patient 60% improvement for speech understanding. This is excellent. But, if a given patient needs an environment with a +10 dB SNR for speech understanding, and he is trying to communicate in a noisy nightclub with an SNR of +2 dB, the 6 dB improvement would only take him to + 8 dB, still 2 dB below the required SNR, and the benefit would be closer to 0% than 60%.

4 In the clinic, a common method to examine a patient’s ability to separate competing signals is to conduct dichotic speech testing. That is, two different speech signals presented with matched timing, and the patient’s task is to repeat back both signals. If the speech signals are very closely matched in time and phase, which can be accomplished with nonsense syllables such as /pa/, /ta/, /ga/, etc., there will be a noticeable advantage for the signals presented to the right ear (because of the superior left brain speech processing). This right ear advantage, however, is not significant for competing sentences (in normal hearing young individuals). A clinical modification of dichotic listening is to have the patient only repeat back the speech signal from a directed ear, the left or the right. For young listeners with normal hearing, and no auditory processing deficit, this is a relatively easy task.

5 Considerable research has been conducted with dichotic listening with older adults, and a significant aging effect has been found. For a relatively simple task, such as repeating back two different sentences (one presented to the right ear, one presented to the left), performance which is typically 95% or higher in young individuals drops to 40-70% in older individuals (poorer performance noted for signal delivered to left ear). Importantly, for these older individuals, 25% continue to score poorly even when they only had to repeat back a sentence for a “directed ear” (either the left or the right). This of course relates directly to group conversations when one attempts to understand a target speech signal from a group of talkers.

6 Two clinical studies have shown that binax provides better than normal hearing in certain demanding environments (University of Northern Colorado, 2014; Oldenburg Hörzentrum, 2013): Speech Reception Thresholds (SRT) in cocktail-party situations improved up to 2.9 dB for wearers with mild to moderate hearing loss using Carat binax or Pure binax hearing aids with narrow directionality, compared to people with normal hearing.