BEYOND AMPLIFICATION: SIGNAL PROCESSING TECHNIQUES FOR IMPROVING SPEECH INTELLIGIBILITY IN NOISE WITH HEARING AIDS

by Brent W. Edwards, Ph.D.
ReSound Corporation
220 Saginaw Drive
Redwood City, CA 94063

INTRODUCTION

The first indication that a person usually has of their hearing loss is difficulty understanding speech in noisy environments such as loud restaurants or crowded rooms. In general, hearing impaired listeners have an abnormally difficult time in such situations even if the speech information is loud enough to be above their threshold of audibility. Several researchers have demonstrated this by showing that hearing impaired subjects require a greater signal-to-noise (SNR) ratio than normally-hearing subjects in order to achieve the same performance in speech-in-noise tests (e.g., Plomp and Mimpen, 1979; Dirks *et al.*, 1982; Pekkarinen *et al.*, 1990).

This deficit is summarized in Figure 1, which replots data from Killion (1997). Using the SIN test, the SNR was adjusted until the listeners understood speech with 50% accuracy. This SNR threshold is plotted as a function of hearing loss. The data shows that as hearing loss increases, the SNR required to achieve 50% performance increases. While someone with no hearing loss can understand speech 50% of the time with a 0 dB SNR, someone with a hearing loss of 60 dB HL requires a 7 dB SNR for the same amount of understanding. So, if both listeners were in the same noisy situation, the person with 60 dB HL would need the speech to be 7 dB higher than the person with 0 dB HL in order for them to understand with the same amount of intelligibility.

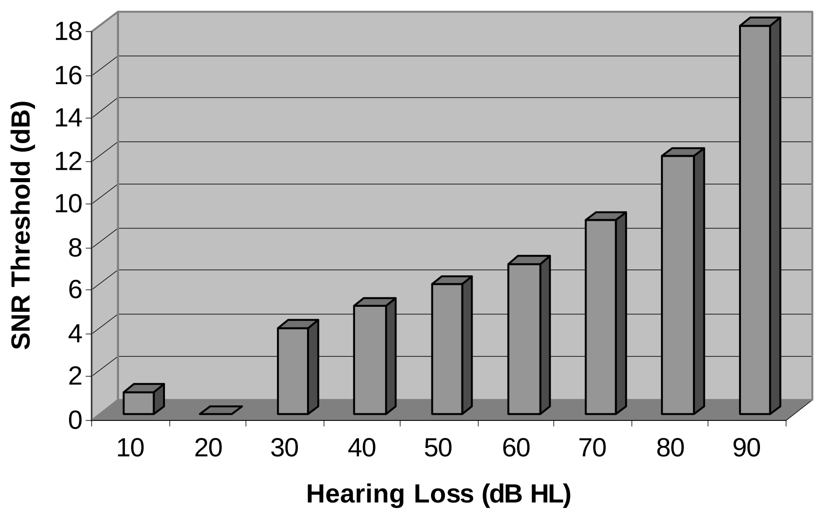

Figure 1*. SNR necessary for 50% word recognition ability relative to normal hearing measured with the SIN Test (from Killion, 1997).*

There are several possible explanations for this deficit that will be discussed later. Since speech understanding in noise remains a major complaint even when hearing aids are worn (Plomp 1978; Tyler *et al.*, 1983; Kochkin, 1994), one of the biggest challenges for hearing aid manufacturers, in addition to restoring audibility and loudness to normal, is to improve speech intelligibility in noise for the hearing aid wearer.

As Killion (1997) has pointed out, hearing aids incorporating technology from over a decade ago were more of a hindrance than a help in noisy situations. Narrow bandwidths, peak clipping, distortion and peaky responses are many of the reasons that hearing aid users did better by removing their hearing aid in adverse environments than by using them. Hearing aid technology has advanced considerably in the past ten years, however, and there are now several solutions to improving the speech-in-noise problem. Directional processing is one obvious technique for improving speech intelligibility in noisy environments. Additionally, the recent introduction of digital signal processing into hearing aids allows more sophisticated techniques to be implemented than could be done with analog technology, although it's important to realize that there is nothing inherent in DSP that improves the sound processing *per se*. The significance of DSP technology is the increased sophistication and complexity of signal processing algorithms that can be implemented. After all, sound processing is ultimately how a hearing aid must be judged, whether analog or digital.

This paper will review several techniques that can improve the understanding of speech in noise. These techniques are not necessarily exclusive; that is, the techniques may be combined to provide more benefit than either on its own. Additionally, techniques that seem as if they should improve intelligibility but do not will also be discussed, along with the reasons that they do not succeed.

DIRECTIONALITY

One of the most obvious and beneficial techniques for improving the intelligibility of speech in noise is the use of directional processing which preserves sound arriving from the front of the hearing aid wearer but attenuates sound arriving from the side and from behind the wearer. Since the speech that a hearing aid user wants to hear is usually in front of them and interfering noise is frequently all around or behind the user, directional-dependent gain can improve the overall speech-to-noise ratio. This is true even if the noise has the same spectral content as the signal and they occur at the same time. Maximum SNR improvement occurs with several microphones that span a distance of several centimeters (Soede *et al.*, 1993; Hoffman *et al.*, 1994). Cosmetic reasons and convenience, however, dictate that microphones be located on the hearing aid body, limiting the number of microphones to two and the distance that they span to less than 15 mm. Given this practical constraint, directionality for hearing aids is implemented with either a single directional microphone or with two omnidirectional microphones placed on the hearing aid. Thompson (1999) provides an excellent comparison of the technical aspects of these two implementations.

Hawkins and Yacullo (1984) found that directional microphones reduced by 3-4 dB the SNR needed for hearing aid wearers to achieve 50% word recognition. According to the data in Figure 1, this would help someone with a hearing loss of 60 dB HL to understand speech in noise as well as someone with a hearing loss of 30 dB HL or less—a significant improvement. Valente *et al.* (1995) found a 7-8 dB improvement in SNR using a hearing aid with a two-microphone array, although this higher level of improvement was most likely due to the placement of the noise source in the direction at which the microphone provided maximal gain reduction—the improvement found when noise surrounds the listener is significantly less. Because of the flexibility of the design, this discussion will focus on directional processing implemented with two omnidirectional microphones.

The basic signal processing structure for directional processing with two omnidirectional microphones is shown in Figure 2. In typical use, these microphones are positioned on a hearing aid such that a line connecting the two microphones points to directly in front and directly behind the hearing aid wearer and is horizontal to the ground. The signal picked up by the rear microphone is delayed by an amount of time similar to the time it takes for sound to travel from one microphone to the other. For a microphone separation of 10mm, that time is approximately 30 microseconds (ignoring near-field effects). The amount of the delay determines the directional pattern (gain as a function of the sound arrival direction). The direction for which the gain is a minimum can be varied by varying the delay. Bachler and Vonlanthen (1997) provide a detailed discussion on this topic.

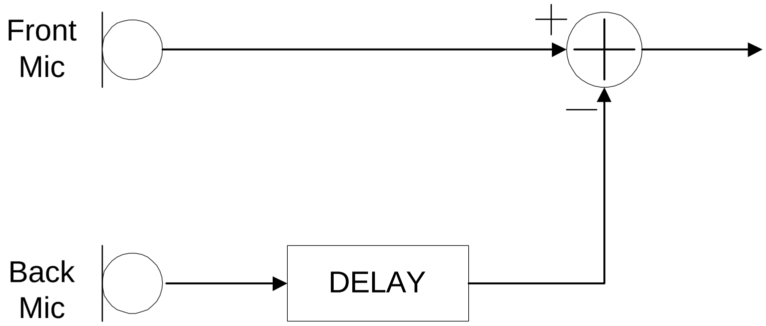

***Figure 2**. Typical configuration for directional processing with two omnidirectional microphones mounted on a hearing aid.*

Typical patterns are shown in Figure 3, where 0 degrees represents the direction directly in front of the wearer while 180 degrees represents the direction directly behind the wearer. The thick line represents the gain as a function of direction of the sound arrival, where the gain from any given direction is represented by the distance from the center of the circle. So, for example, the pattern on the right (cardioid) attenuates sound the most from directly behind the wearer, while the pattern on the left (bidirectional) attenuates sound the most from directly to the left and to the right of the wearer.

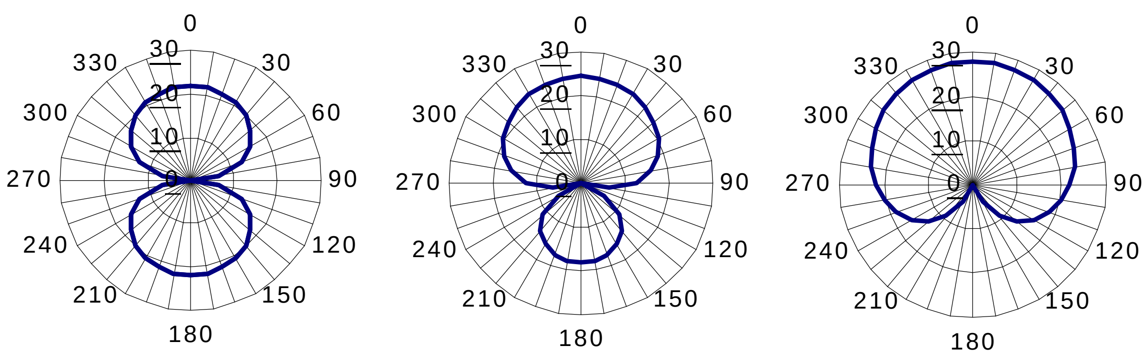

***Figure 3.** Three typical directional patterns that can be achieved with the processing shown in Figure 2. From left to right, the patterns are: bidirectional, hypercardioid, cardioid.*

Assuming that the target speech is directly in front of the listener, the improvement to SNR provided by a specific directional pattern depends on the direction or directions of the noise sources. If the noise only arrives from the same direction as the target speech, then there is no improvement to the SNR with directionality. If the noise source is directly behind the user, then the cardioid pattern provides the greatest improvement. Typically, however, interfering sound sources in noisy environments arrive at the listener from all directions, particularly in reverberant environments, in which case the pattern in the middle (hypercardioid) provides the greatest SNR improvement.

The directivity index (DI) is common measure of directional performance first introduced to quantify antenna array performance (Uzkov, 1946) and is the ratio of the gain from 0 degrees (directly in front) to the average gain from all directions. Improvements to the speech reception threshold (SRT) provided by directionality when measured with diffuse noise, such as with the HINT test, are similar to the DI value in dB. In other words, a hearing aid with an average DI of 4 dB would improve the SRT by about 4 dB. Under conditions where the noise is diffuse or has equal intensities from all directions, the hypercardioid pattern in the middle of Figure 3 has a DI of 6 dB when the hearing aid is placed in the free field. This is the maximum directionality achievable by a two-microphone array under diffuse noise conditions (Kinsler and Frey, 1962). The other two patterns shown have a DI of 4.8 dB in free field.

Recent modifications to the DI measure incorporate the frequency importance function of the Articulation index (AI) to take into account the frequency-dependent characteristic of directionality (e.g., Peterson 1989; Greenberg and Zurek, 1992). This modified measure is known as the AI-DI (Roberts and Schulein, 1997) and better characterizes the effect of directionality on speech intelligibility.

Wolf *et al.* (1999) demonstrated that hearing aid wearers prefer that directionality not be on all of the time—users desire the ability to switch between directional and omnidirectional modes. One Improvement that two omnidirectional microphones provide over one directional microphone is that directionality can be switched on and off by turning the back microphone on and off. Because one pattern may be better for situations where noise is coming from directly behind the hearing aid wearer but another pattern may be better for situations where noise is coming from all directions, a further improvement to directional hearing aids is to allow the pattern to be selectable for the specific need of the wearer. This could be implemented by either allowing the pattern to be programmed when the hearing aid is fit or by allowing the wearer to switch from one pattern to another.

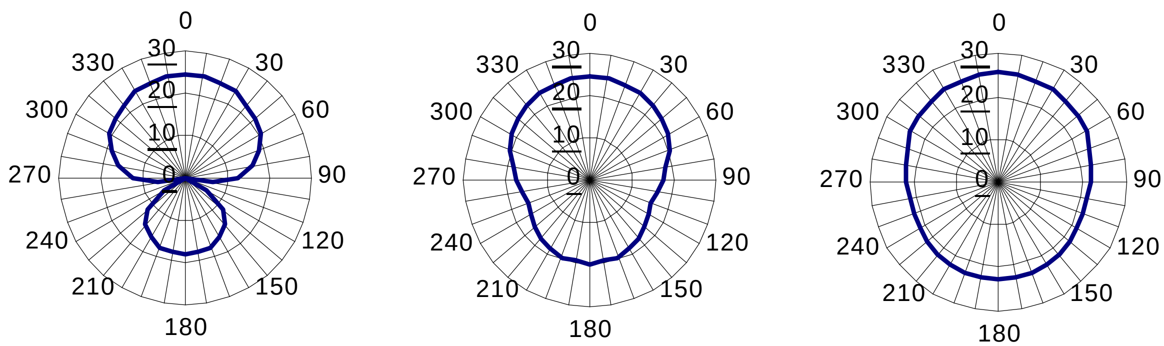

Figure 4. *The effect on directionality of a sensitivity mismatch between the two omnidirectional microphones. From left to right, the mismatches are: 0 dB, 1 dB, 2 dB. The DI at 1kHz for these cases are 6dB, 4.5 dB and 2.7 dB, respectively. See text for details.*

ReSound's Digital 5000 series allows both possibilities by implementing the delay digitally. This allows different directional patterns to be programmed such that the optimal directional pattern, and thus optimal SNR improvement, can be selected by the wearer for the given noisy environment. A different directional pattern can also be software-programmed into the device depending on whether the users tend to have trouble understanding speech when noise is primarily behind the user or all around the user.

One difficulty with implementing two-microphone directionality is that the microphones must be very closely matched at all frequencies of interest in order to provide good SNR improvement at these frequencies. If the frequency responses of the two microphones are not identical, then directionality is degraded (Thompson, 1999). This is illustrated in Figure 4, taken from Edwards *et al.* (1998b). The figure on the left shows a hypercardioid pattern obtained when microphones are perfectly matched, resulting in a DI of 6 dB. The figure in the middle shows how directionality is degraded when one microphone is 1-dB less sensitive than the other microphone at 1 kHz with a microphone separation of 12 mm. The noise from behind the listener in this condition is not as significantly attenuated and the resulting DI at 1 kHz is 4.5 dB. In the figure on the right, a 2-dB difference exists between the sensitivities of the two microphones and the directional pattern appears to be nearly omnidirectional, with a DI of 2.7 dB.

Since the effect of sensitivity mismatch on the DI varies with frequency, the reduction in DI across all frequencies must be calculated in order to determine the effect of a mismatch on speech intelligibility. Figure 5 plots the AI-DI of a directional device as a function of the sensitivity mismatch between microphones where the mismatch is identical across all frequencies. This was calculated analytically for both a hypercardioid and cardioid pattern in a free field. The spacing between the microphones was assumed to be 12 mm. The data again shows that performance decreases as the microphone mismatch increases. What is interesting to note is that even though the hypercardioid is more sensitive to microphone mismatches than the cardioid (note the steeper drop in performance as the microphone mismatch increases), the hypercardioid still outperforms the cardioid for microphone mismatches of 3.5 dB or less. In fact, the directional benefit provided from a cardioid with perfectly matched microphones is the same as the benefit provided by a hypercardioid pattern with a mismatch of 1.5 dB.

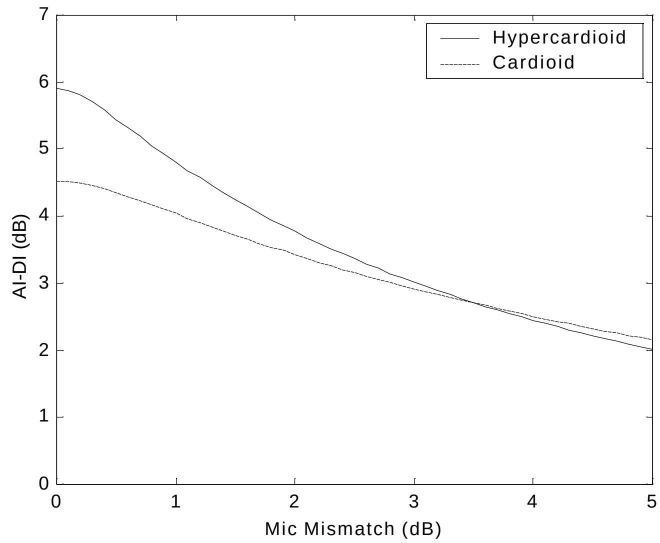

Figure 5. *Effect of mismatched microphone sensitivity on the AI-DI, showing that the AI-DI is reduced as the mismatch is increased. The solid line shows the effect on a hypercardioid pattern and the dashed line shows the effect on a cardioid pattern.*

While these figures are patterns obtained in the free field, similar reductions in directional benefit will also be obtained when worn on the head, resulting in poorer speech intelligibility for microphones that are not perfectly matched. Thompson (1999) speculated that this problem may eventually be solved by compensating for the sensitivity differences with digital signal processing, as long as sufficient DSP power exists and the signal from each microphone is digitized with a separate A/D. This solution does indeed exist already, and is implemented in ReSound's Digital 5000 by digitizing each microphone signal with separate A/Ds and implementing a digital filter to match the responses of the two microphones. Similar to the manner in which Thompson describes, the sensitivities of the microphones are measured and the parameters of the digital filter are set such that the microphone match is improved and the directional benefit optimized. A necessary feature to allow a digital hearing aid to implement all of these advances is that the hearing aid have 2 A/Ds; otherwise, the microphones can't be matched with a digital filter and the delay can't be digitally programmed.

Thus, digital processing is used in ReSound's Digital 5000 to benefit directionality, and thereby improve speech intelligibility, in two ways: (i) by implementing the delay to the back microphone digitally, the directional pattern can be selected to optimize the SNR improvement for a given noisy situation, and (ii) by implementing a digital filter which matches the sensitivity between the two microphones in order to maximize the AI-DI of the device.

NOISE REDUCTION

The term noise reduction will be used here to describe techniques that process a single signal which contains the noise and speech together (directional techniques process at least two signals and use differences in those signals to reduce the noise). Separating the desired speech from the competing noise after they've both been picked up by the same microphone is extremely difficult, particularly when the competing noise is speech from one or more other talkers.

Ideally, a noise reduction algorithm will take advantage of temporal and/or spectral mismatches between the target speech and the interfering noise in order to remove the interference while preserving the target. Unfortunately, it is difficult for a noise reduction algorithm to identify which part of the signal is the target speech and which part of the signal is the interfering noise since both signals usually have energy that occupies the same spectral and temporal coordinates. Even if the noise and target speech never have the same frequency components at the same time so that the gain can be reduced on the noise signal without affecting the speech signal, the noise reduction algorithm would still have to figure out which spectro-temporal regions were speech and which were noise. This is difficult because many noises have speech-like characteristics and many parts of speech have noise-like characteristics.

Lim (1983) has noted that there exist many techniques for improving the SNR of speech in noise, some by as much as 12 dB, but that none of them improve intelligibility. One reason is that improving the *acoustic* SNR does not necessarily improve the *perceptual* SNR. The human speech recognition system, which includes the auditory system, has its own sophisticated "noise reduction" techniques which allow humans to understand speech in the presence of a wide variety of acoustic interference. This system no doubt analyzes the many spectro-temporal cues that exist in the degraded speech signal to produce meaning. This situation is worsened by hearing loss where the damaged cochlea distorts the signal and the speech cues being received by the auditory cortex. Thus, one way that a hearing aid could improve the *perceptual* SNR would be to alter the acoustic signal in order to restore to normal the *perceptual* signal. This normalization technique differs from noise reduction and will be discussed in a later section. In this section, techniques that do not explicitly take into account a listener's hearing loss will be discussed.

Spectral subtraction techniques (Boll,1979), common in such fields as the telecommunications industry, attempt to estimate the amplitude spectrum of noise and subtract this from the amplitude spectrum of the speech and noise combined. This technique has had little success in improving intelligibility, partly because of the difficulty in estimating the noise spectrum separately from the speech spectrum. Levitt *et al.* (1986) evaluated the original implementation of this algorithm (Weiss *et al.*, 1974) with hearing-impaired subjects but had limited success. Even though the algorithm significantly improved the signal's SNR, an examination of the "noise reduced" signal revealed that the vowel formants were enhanced but the noise-like consonants were removed. One additional problem with these spectral subtraction techniques is the side effect of audible distortion, known as "musical noise" because of the noise's tone-like quality. Whether this audible distortion would be acceptable to a hearing aid wearer is unknown.

Several algorithms have been developed which attempt to reduce the gain in frequency regions where noise occurs, and initial investigations into this gain reduction technique several years ago appeared to result in improved intelligibility (e.g., Stein and Dempsey-Hart, 1984). Devices developed in the late 1980's known as ASP devices implemented this idea with so-called BILL (bass increase at low levels) processing. Lybarger (1947) first proposed this approach which is based on the theory that high-level, low-frequency sounds can mask low-level, high-frequency sounds. Since most environmental noises have much of their energy in the low frequencies (Klumpp and Webster, 1963) and the information carrying consonants of speech are in the high frequencies, it seemed that upward spread of masking of the noise could make low-level consonants inaudible. Thus, it seemed that reducing the gain in the low-frequency region when the level in this region was high would prevent high-frequency speech from being masked and thereby improve intelligibility. This technique is in fact used in some digital hearing aids currently on the market.

Despite the seemingly obvious reasons why this technique should succeed, research has not found any benefit from this technique for speech intelligibility (Punch and Beck, 1986; Neuman and Schwander, 1987; Van Tasell et al, 1988; Tyler and Kuk, 1989; Fabry and Van Tasell, 1990; Van Tasell and Crain, 1992). Reducing the gain in the low frequency region, in fact, risks making low-frequency speech cues inaudible. Studies which did show improvements to intelligibility when the gain in regions of narrowband noise was reduced (Festen et al, 1993; Rankovic, 1999) used extremely high levels of noise unlikely to be found in real situations. Levitt (1991) noted that reducing the gain in regions of noise is unlikely to improve intelligibility in practical use of a hearing aid because upward spread of masking is typically negligible at realistic level of noise experienced by hearing aid wearers, although he was careful to note that there were some individuals in previous studies which did show some improvement.

Because of the limited success of improving intelligibility using single-microphone noise reduction techniques, an alternate goal for such algorithms is to improve the sound quality of speech in noise, thereby improving the "comfort" of the wearer. While certainly a less impressive goal, this can still have significant impact on a hearing aid wearer who is attempting to listen to speech in a noisy environment over a prolonged period of time. It is possible that listening to speech in a less noisy background may require less effort by the listener even though it doesn't increase the overall understanding of speech, providing for a less exhausting experience. Indeed, less effort required of the hearing aid wearer may allow them to perform dual-attention tasks that normal listeners typically take for granted, such as listening to a talker while also paying attention to their surroundings, simultaneously noting other conversations that may be going on around them or monitoring the activity of other people in the room. Simply eliminating a distracting or annoying sound, such as a nearby air conditioner, could also provide significant benefit to a hearing aid user because of the improved sound quality. In support of this, Kochkin (1999) has found that the second most common reason that owners don't use their aids is because of background noise complaints. Thus, reducing noise while preserving speech intelligibility is a reasonable and useful signal processing goal for a noise reduction algorithm.

ReSound's Digital 5000 noise reduction algorithm addresses this issue in a number of ways. The signal is spectrally segmented using 14 bandpass filters, and a speech-to-noise ratio is estimated in each. This is done by measuring the amount of modulation, or level fluctuation, in each band. The level of speech in quiet fluctuates over a 30-dB range and has modulation frequencies centered on 4Hz (Houghast and Steeneken, 1985), the average occurrence rate of syllables. Environmental noise typically has much lower modulations than speech, even if the noise is multi-talker babble. Figure 6 shows the output of an octave filter centered on 2 kHz for speech in quiet and for 6-talker speech babble. The plots clearly show that the amount of modulation in the speech-alone signal is significantly greater than the modulation in the babble noise.

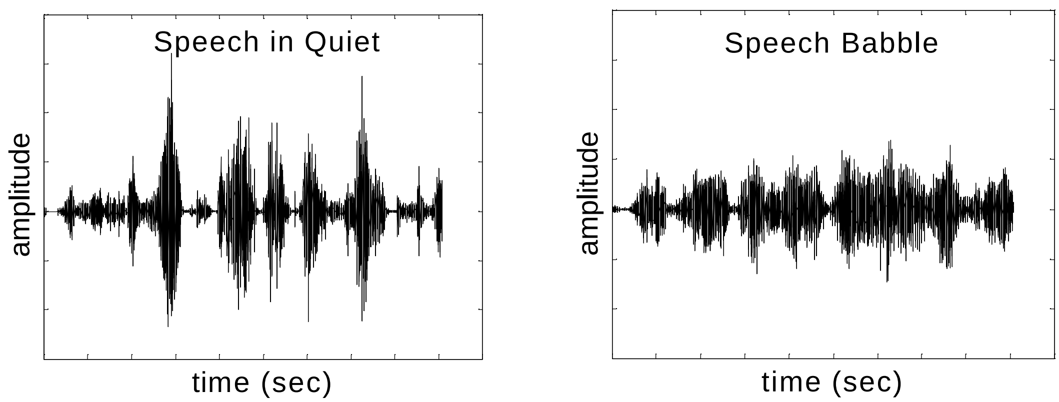

Figure 6. *Speech in quiet (left) and 6-talker babble (right), both filtered through an octave band at 2kHz, demonstrating that speech in quiet has higher level fluctuations or envelope modulation than speech babble.*

Additionally, as noise is added to speech, the modulation of the combined speech and noise is reduced. Because of this effect, the amount of speech in any given band can be estimated by measuring the amount of modulation in that band. This is demonstrated in Figure 7, which replots the data from Edwards *et al.* (1998b), showing the general trend for modulation to decrease as the level of speech decreases relative to the level of noise. The algorithm in ReSound's Digital 5000 decreases the gain in each band as the amount of modulation decreases in a band, making the noise in that band less audible.

As stated previously, it is important that the noise reduction processing not reduce speech intelligibility while improving sound quality. Because of the spectral resolution provided by the 14-band filterbank, the frequency regions with low levels of modulation can be isolated from those with high levels of modulation, and the gain reduction can be better contained to those frequency regions with noise. Similar techniques that use fewer, wider filters risk reducing the gain in spectral regions where noise does not exist but which fall within the filter bandwidth that has its gain reduced. If, for example, noise has no spectral components above 500 Hz, a broad lowpass filter that extends up to 800 Hz would reduce the gain up to 800 Hz, unnecessarily reducing the level of speech from 500 to 800 Hz. These noise reduction techniques with wide filters would then be more likely to adversely affect intelligibility by unnecessarily reducing the audibility of speech. Implementing noise reduction with 14 bands in ReSound's Digital 5000 reduces this possibility.

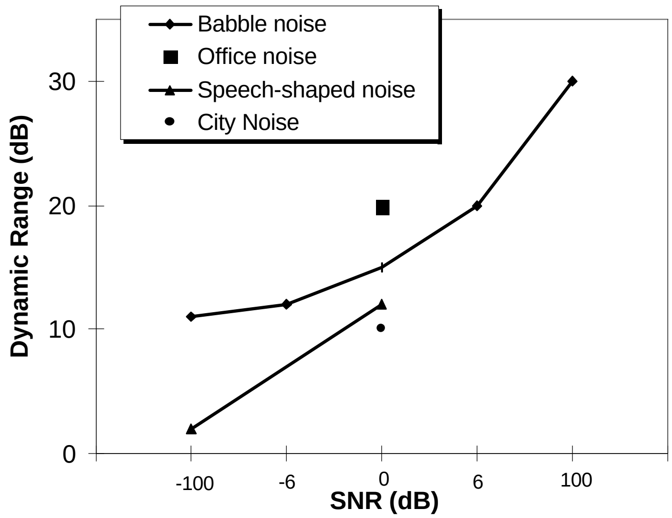

Figure 7. *The dynamic range of speech added to different types of noise, showing that the dynamic range is reduced as the SNR decreases.*

FEEDBACK CANCELLATION

Feedback is a significant problem for many hearing aid users, since it not only limits the gain that the hearing aid can provide but is an annoyance to the hearing aid wearer and those around them. Solutions to this problem usually deteriorate speech intelligibility since most solutions reduce the gain in frequency regions where feedback occurs, limiting audibility of speech at those frequencies.

Feedback exists when sound from the output of a hearing aid leaks back to the microphone. This is not a problem unless the gain from the microphone to the receiver exceeds the attenuation from the receiver back to the microphone (both measured in dB). Consider the case where the level of the hearing aid output at 3kHz is 80 dB SPL and 60 dB SPL of that sound leaks back through the vent, representing an attenuation from the receiver to the microphone at 3kHz of 20 dB. Feedback oscillation will occur with this system if more than 20 dB of gain is provided by the hearing aid at 3 kHz.

This suggests two solutions to reducing the feedback problem: (i) increase the attenuation from the receiver to the microphone, or (ii) reduce the gain from the microphone to the receiver. The first strategy can be achieved by either reducing the vent size or closing the vent altogether. This, however, is typically not done except in severe cases because reducing the vent size increases the patient's feeling of occlusion which is objectionable. Therefore, the second strategy is usually followed, which is to limit the amount of gain that the hearing aid provides in frequency regions of feedback problems. Unfortunately, feedback typically occurs at high frequencies which is exactly where patients usually need the most gain since the majority of hearing aid wearers have high-frequency loss.

There exists another strategy, implemented in ReSound's Digital 5000, which effectively increases the attenuation from the receiver to the microphone without actually changing the acoustics of the feedback path. Ideally, if the DSP in the hearing aid knew exactly what the signal was that was feeding back from the receiver to the microphone, then the DSP could generate exactly the same signal and subtract it from the incoming microphone signal. In this case, the feedback signal is cancelled in the DSP. Since only the feedback signal is effected, any other signals picked up by the microphone from the acoustic environment is unaffected by this processing and can be amplified by the hearing aid in a normal manner.

Ideally, this technique would completely cancel the feedback signal such that no limit exists on the gain that could be provided by the hearing aid. In a practical implementation, the feedback signal cannot be perfectly predicted by the DSP and thus the feedback signal cannot be completely eliminated. A reasonable estimate of the feedback signal can be calculated by the DSP, however, such that the feedback signal is significantly reduced, effectively increasing the feedback attenuation. The result of this is that the amount of gain that can be provided by the hearing aid is increased by the amount that the feedback signal is reduced (both measured in dB).

Clinical testing on over 20 subjects has found that the majority of patients experience an increase of 6-10 dB in useable gain with the activation of the feedback cancellation algorithm. So, if the feedback signal is reduced by 10 dB with this algorithm, the maximum amount of gain that can be provided by the hearing aid without feedback is increased by 10 dB. This means that more gain can be provided by the hearing aid with the feedback cancellation algorithm active than without it, resulting in increased audibility if the patient's audiogram required more gain than could be provided without the algorithm. In the previous example where the gain at 3 kHz was limited to 20 dB, this algorithm would allow up to 30 dB of gain.

Figure 8 demonstrates the increased gain obtained with a subject that had a severe hearing loss. The dashed line shows the gain that could be programmed, measured in a 2cc coupler, where the gain in the high frequencies was below her prescription because of feedback problems. The solid line shows the gain that was provided once the feedback cancellation algorithm was activated. Approximately 7 dB more gain was provided in the high-frequency region because the feedback signal was cancelled and more gain could be programmed. This subject's HINT in quiet score increased by 4.3 dB with the feedback cancellation algorithm active and the added gain.

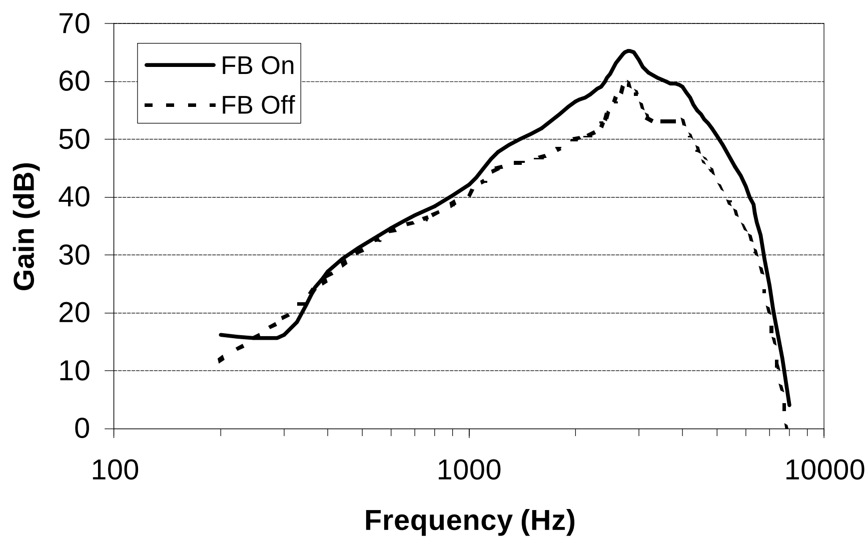

Figure 8. *2cc coupler measurements showing the increased gain that can be achieved when the feedback cancellation algorithm is active. Both curves have the high frequency gain set to the maximum level before the onset of feedback. The dashed curve shows the gain when the algorithm is off, and the solid curve shows the gain when the algorithm is on.*

According to the articulation index, increasing the audibility of speech has a direct and calculable result on speech intelligibility. While it is clear that more gain will increase speech audibility (and intelligibility) in quiet for those who need it, more gain in environments where background noise limits speech audibility may not necessarily improve intelligibility since the added gain will amplify both the noise and speech by the same amount. Thus in noisy situations, the articulation index may not be increased since the SNR at any given frequency is not improved. There is a situation where increasing the gain in noisy environments can improve intelligibility, however. This is the case where the audibility of high frequency speech cues is limited not by the noise spectrum but by the hearing loss of the listener. Speech cues may exist in high frequency regions, even with noise in those regions, but the listener may not be able to hear them because their hearing loss exceeds the level of both the speech and noise sounds. By providing more gain in these regions, previously inaudible speech cues may be made audible. Indeed, French-St. George *et al*. (1993) found that speech understanding improved in speech babble using feedback cancellation in a prototype digital device.

The key to the feedback cancellation algorithm implemented in ReSound's Digital 5000 is the proper mimicking of the acoustic feedback path in the DSP. If the feedback path can be correctly modeled, then the signal at the receiver can be run through this model within the DSP to predict the feedback signal that will be generated acoustically. Figure 9 shows how the DSP obtains an estimate of the feedback path from the receiver to the microphone. During the fitting of this hearing device, a noise is generated within the DSP while the throughput from the microphone to the receiver (i.e., the compressor) is turned off. The DSP measures how much of the noise from the receiver returns to the microphone and then calculates the acoustic properties of the feedback path.

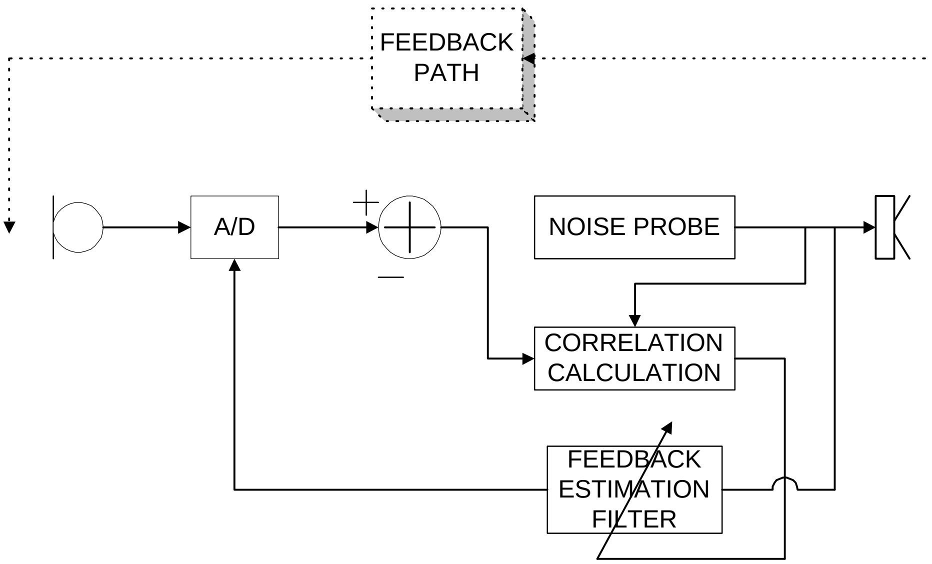

Figure 9. *A schematic showing how the characteristics of the acoustic feedback path is estimated. The dotted line represents the acoustic feedback path through a vent from the receiver to the microphone.*

A digital filter is then designed which mimics the response of the acoustic feedback path so that when the receiver signal is fed through the digital filter, the output of the digital filter is similar to the actual acoustic feedback signal that will be picked up by the microphone. As shown in Figure 10, the DSP then simply subtracts the estimated feedback signal from the microphone input, subtracting out the feedback signal but not affecting any other signals being picked up by the microphone such as speech.

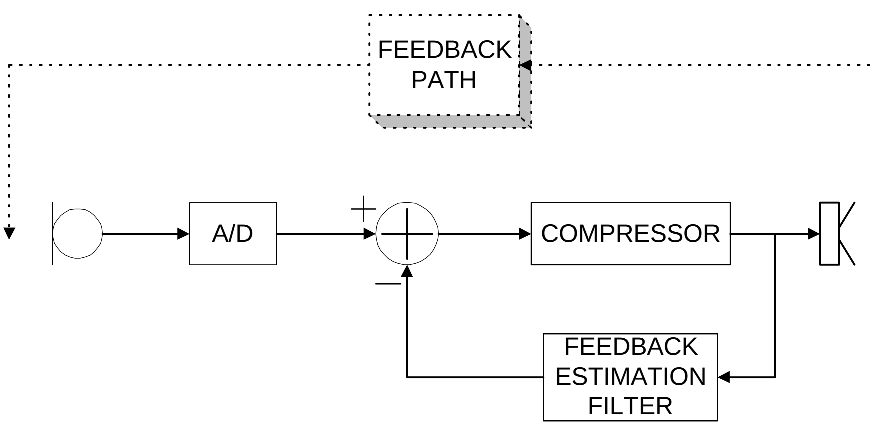

Figure 10. *Schematic of how the feedback algorithm is incorporated in the hearing aid processing. The Feedback Estimation Filter should have the same transfer function as the Feedback Path.*

So, the feedback cancellation algorithm in ReSound's Digital 5000 not only reduces the annoyance of feedback problems for hearing aid users, allowing them to wear more open vents, but also improves speech understanding for both quiet and noisy conditions.

HEARING LOSS COMPENSATION

As stated in the beginning of this article, something about the nature of hearing loss detrimentally affects a person's ability to understand speech in noise beyond simple loss of audibility. All of the techniques for improving the intelligibility of speech discussed so far, however, have been unrelated to the primary cause of the hearing aid wearer's speech understanding deficit: the patient's hearing loss. While the following statement seems tautological, it bares stating simply to emphasize what is obvious but often overlooked: *speech understanding* would be restored to normal if *auditory perception* were restored to normal. Since sensorineural hearing loss detrimentally affects speech intelligibility in noise, one strategy for improving intelligibility is to process sound such that the listener's cochlear damage is counteracted. While the feasibility of achieving this goal remains unclear, it is a goal that can drive many strategies for improving hearing and at the same time improve speech understanding.

The processing performed on acoustic stimuli by the auditory periphery, from the pinna to the auditory nerve, is a well defined process. If any part of the auditory peripheral system does not function properly, then the signal transmitted to the brain by the auditory nerve will be distorted; the brain will not receive the same signal that it would have if the auditory periphery were functioning properly. The majority of people's hearing loss is due to outer hair cell damage, so the acoustic signal received by the cochlea is not properly transmitted to the auditory nerve, resulting in the auditory nerve transmitting a distorted signal to the brain. The objective of a hearing aid should be to process the signal such that the signal received by the auditory nerve, and thus the brain, is restored to normal as much as possible. If inner hair cell damage exists, then this becomes impossible since the communication link to the auditory nerve fibers that innervate the dead inner hair cells is severed and the normalization strategy must be revised. This problem with inner hair cell damage will be discussed later in this section.

How do we know the way in which the auditory nerve signal is distorted so that a hearing aid can be designed to counter the distortion effect? The most direct way would be to measure a person's auditory nerve response to different stimuli, but for obvious reasons this is not an option. A less direct approach is to use psychological techniques whereby sounds are played to a person with hearing loss and then questions are asked about the sound. The simplest of these techniques is in fact what is done to obtain a person's audiogram, where the person's sensitivity to soft sounds is measured and compared to the sensitivity of normally-hearing people. More sophisticated techniques have been used recently to obtain more detail about how damaged outer hair cells distort sound.

One technique that has become a staple of perceptual characterization is loudness growth measures which demonstrate that hearing impaired listeners show loudness recruitment, whereby the reduction in their sensitivity to sound is greatest at low levels and sensitivity becomes more normal as the level increases. This has led to the design of compression in hearing aids which reduces the gain as signal levels increase in order to restore the loudness of sounds to normal. In terms of our stated goal for hearing aids, compression in a hearing aid attempts to bring the level of the signal at the auditory nerve to normal. The rate-level functions, or input-output functions, of auditory nerve fibers have been known to be steeper in animals with damaged cochleas, and this phenomenon is directly related to loudness recruitment observed psychoacoustically. By compressing sounds in the hearing aid, the psychological goal is to restore loudness perception to normal and physiological goal is to restore the rate-level functions of auditory nerve fibers to normal.

While loudness recruitment has been know to be a consequence of hearing loss for several decades (Fowler, 1936), it is only recently that recruitment has been shown to be a consequence of outer hair cell damage (Ruggero and Rich, 1991). More specifically, Ruggero and Rich discovered that healthy cochleas naturally compress sound as a part of its normal auditory function and that damage to outer hair cells eliminates compression in the cochlea. Figure 11 replots their data, where the solid line shows the response of the basilar membrane in a healthy cochlea as a function of stimulus level and the dashed line shows the response in a cochlea with deadened outer hair cells. If a hearing aid were designed with the intent to restore the response of the basilar membrane to normal, one of the characteristics of the sound processing would have to be a compression function which performs the compression that the damaged cochea can no longer do. It is fairly straightforward to see how the phenomenon of loudness recruitment is a result of this loss of compression in the cochlea, a loss caused by damaged outer hair cells.

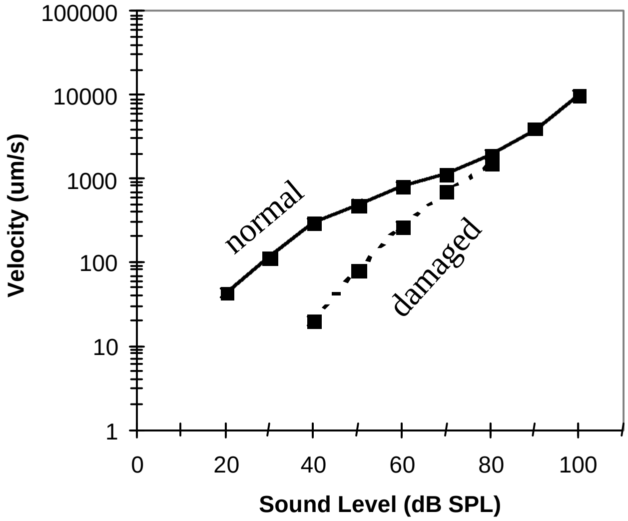

Figure 11. *The response of the basilar membrane in a healthy colchea (solid line) and one in which the outer hair cells were deadened (dashed line). Note that the healthy cochlea compresses sound but that this compression is lost in the damaged cochea. (Replotted from Ruggero and Rich, 1991.)*

The fact that a healthy cochlea compresses sound and a damaged cochlea does not compress sound in regions of outer hair cell loss is a fact which has not been fully appreciated by some. Several researchers (e.g., Plomp, 1988) have noted that compression hearing aids distort sound by reducing envelope fluctuations, as shown in the upper half of Figure 12.

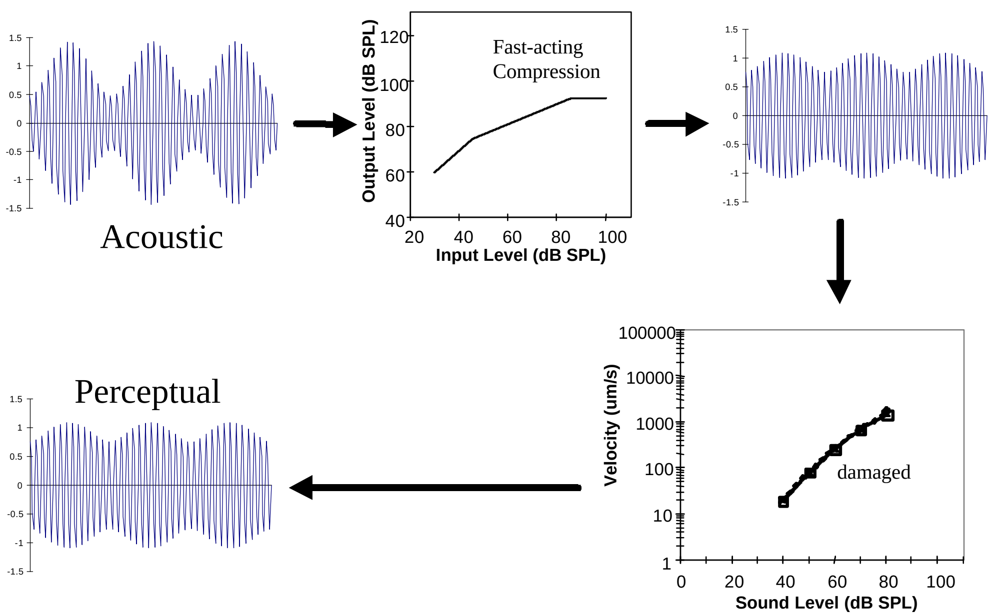

Figure 12. The effect of fast-acting compression on the perception of a signal's envelope. The upper box represents the hearing aid's input-output function and the lower box represents the cochlea's input-output function. While the cochlea no longer compresses the envelope because of outer hair cell damage, the envelope of the signal is still perceived normally (compare to Figure 13) because of the compression performed in the hearing aid.

It has been suggested that this obvious distortion of sound could result in at least abnormal sound perception and at worse poorer speech intelligibility. This last claim is based on the fact that the envelope of speech carries certain cues used for speech understanding (Erber, 1972; Van Tasell *et al.*, 1987b; Shannon *et al.*, 1995), such as differentiating plosives from fricatives (Dreschler, 1989). While the temporal fluctuations of speech do indeed carry important consonant information as well as information about phonemic and syllabic boundaries, the claim that compression is detrimental to someone with hearing loss is similar to the claim that eyeglasses are detrimental to written word recognition for the visually impaired because eyeglasses distort visual cues. The effect of the audio/visual aid on the *physical* signal is not the significant factor, what is important is the effect of the aid on the *perceptual* signal. In the case of hearing aids, the effect of the hearing aid processing must be considered in combination with the effect of cochlear processing, as shown in Figure 12. Here, it is shown that the effect of hearing aid compression and damaged cochlear processing is similar to the effect of healthy cochlear compression shown in Figure 13. If the hearing aid had not compressed the signal, the perceptual signal after the damaged cochlear processing would be abnormally expanded—the envelope fluctuations would be exaggerated relative to normal perception. This concept was proven experimentally by Wojtczak (1996) and Moore *et al.* (1996).

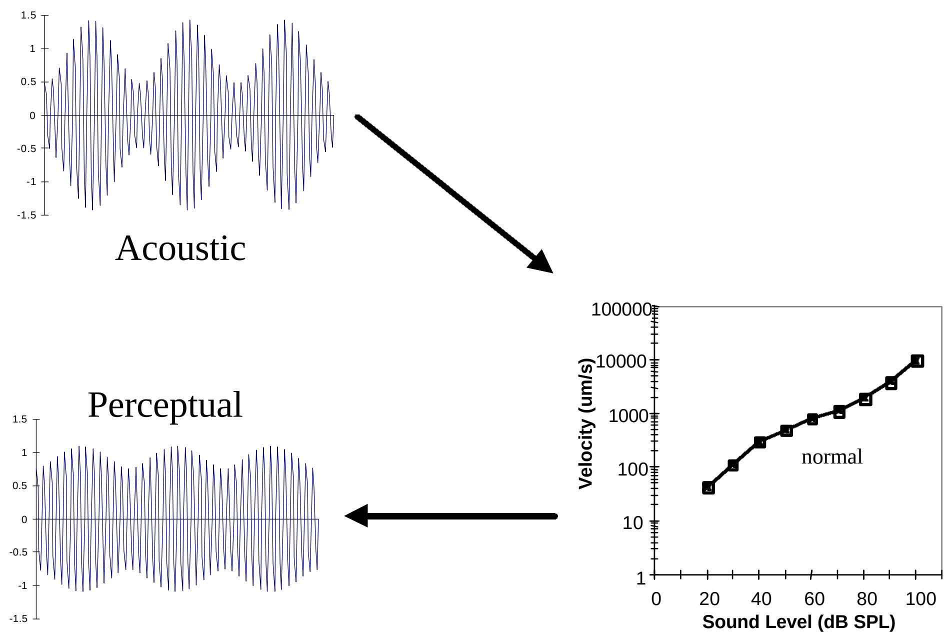

Figure 13. *The effect of a healthy cochlea on the perception of a signal's envelope. Because of compression in the cochlea, the envelope of the perceptual signal has been compressed relative to the acoustic signal. This compression is lost with the majority of hearing loss*

Over the past decade, scientists have begun to realize that the consequences of lost physiological compression extend beyond raised thresholds and loudness recruitment. Several of these are well summarized by Moore and Oxenham (1998), and are briefly noted here. Recent research has shown that outer hair cell damage has the following detrimental results on auditory perception: increased off-frequency masking (Nelson and Schroder, 1996); poorer frequency resolution (Van Tasell *et al.*, 1987a); poorer temporal resolution (Buus and Florentine, 1985); increased forward and backward masking (Moore and Oxenham, 1998); poorer temporal integration (Oxenham *et al.*, 1997); distorted envelope perception (Wojtczak, 1996). All of these effects distort the signal received by the auditory nerve and thus distort the perception of sound relative to how it would be perceived if the cochlea were healthy. Given all of the other effects that result from loss of compression, it is not surprising that restoration of speech in noise to audible levels still results in an intelligibility deficit for the majority of people with sensorineural hearing loss.

As has been suggested, one strategy for improving speech intelligibility in noise is to design a hearing aid which restores all of these perceptual consequences of hearing loss to normal. For example, subjects with moderate hearing loss have an impaired ability to detect gaps in noise (Buus and Florentine, 1985), which can affect the ability to identify consonants. Glasberg and Moore (1992) have suggested that compression can improve the ability of someone with hearing loss to detect temporal gaps, and indeed Dreschler (1989) found that compression improved the ability of hearing impaired listeners to discriminate plosives from fricatives which he attributed to their improved ability to detect the gap which precedes the burst in plosives.

Obviously, in order to compensate for processing that is lost due to a damaged auditory system, a hearing aid would have to somehow provide the processing that is lost. In the case of damaged outer hair cells, that processing is compression which is fast-acting (Yates, 1995) and frequency-specific (Ruggero,1992). While compression in two or three bands is a good start to this attempt, it is a far way from providing the compression that is done by a healthy cochlea. First of all, compression is localized in frequency. The filters that are associated with compression are not two or three filters contiguous in frequency, but are thousands of filters that have significant overlap. If every inner hair cell is considered an auditory filter, then each filter has a center frequency that is hundredths of an octave separated from adjacent filters yet has a bandwidth that is closer to a third of an octave. This provides a significant overlap in the filters and has several beneficial consequences in how the signal is transmitted to the auditory nerve. A similar overlapping filter design for hearing aids, derived by Lindemann (1997) from the perspective of eliminating distortion in the filter transition regions of traditional filter banks (Edwards and Struck, 1996), is the basis for the design of the compressor in ReSound's Digital 5000. With significantly more filter bands than previously implemented in hearing aids, with a heavily overlapping filterbank structure, and with bandwidths that are more similar to those of auditory filters, this compression architecture is the next generation in multiband compression and a necessary step towards providing the compression lost due to sensorineural hearing loss. This 14-band wide dynamic range compressor, shown in Figure 14, is described in more detail in Edwards *et al.* (1998a).

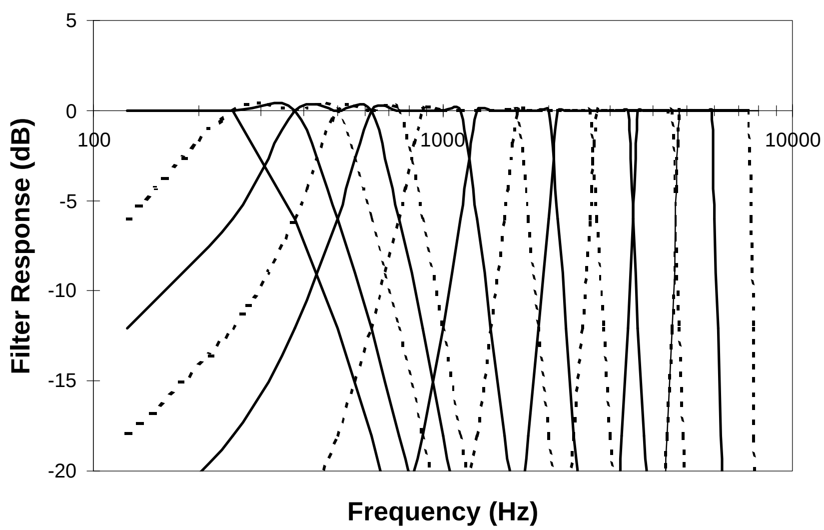

Figure 14. *The 14 filters used in ReSound's Digital 5000 comprFigure 14. The 14 filters used in ReSound's Digital 5000 compressor. Because of the heavily overlapping nature of the filter structure, individual filters are difficult to distinguish.*

Of course, the cochlea is a highly sophisticated transducer that decomposes the acoustic signal into approximately 3500 subbands with significant nonlinearities involved. It may be that a hearing aid will never be able to perfectly reproduce the nonlinearities lost with dead outer hair cells. It is clear, however, that hearing aids have much farther to go before they reach the limit of what they can do to restore perception to normal. A much more difficult problem exists with hearing loss caused by inner hair cell damage. In this case, no amount of amplification will restore the signal to normal in the auditory nerve fibers that innervate the dead inner hair cells—the transmission lines to those fibers have been effectively severed. If there is enough inner hair cell damage along a region of the basilar membrane (corresponding to a spectral region), then another strategy may need to be taken for the hearing aid processing. In this case, the strategy may be to maximize the transfer of speech information to the brain through whatever means possible. An extreme example of this type of strategy for the visual system is Braille. Written words can no longer be trasmitted through the normal neural mechanisms to the brain, and thus a different strategy for information transfer is taken. In the case of Braille and the visual system, a completely different sensory mechanism is used. While vibro-tactile aids also stimulate a different sensory mechanism for people with very little residual hearing, people with less severe inner hair cell damage may benefit from a less drastic remapping of the speech information. One recent example that has shown promise is frequency shifting, where the spectral content of speech is shifted down to lower frequency regions where listeners still have working inner hair cells to transmit the speech information (Turner and Hurtig, 1999). Other strategies may exist where the processing done by the hearing aid is not intended to restore perception to normal but to maximize the transfer of speech information through the reduced capacity of the damaged auditory system.

CONCLUSIONS

Improving the intelligibility of speech in noise is one of the most difficult tasks facing hearing aid manufacturers today. There are currently a variety of tools available for this task, in part due to the application of digital signal processing to hearing aids. With DSP, more sophisticated signal processing algorithms can be implemented than were achievable before, as evidenced by the following design features provided by ReSound's Digital 5000:

- selectable directional patterns achievable by implementing the delay to the rear microphone digitally, which allows the selection of the directional pattern that maximizes the SNR improvement for a given noisy environment;
- digital matching of the two omnidirectional microphones after both microphone signals are digitized with 2 A/Ds, which maximizes the directivity index by minimizing the microphone mismatch;
- feedback cancellation that increases the gain limit at which unstable feedback occurs, which increases the audibility of speech in many situations;
- 14-band noise reduction, which improves the sound quality of speech in noise;
- a 14-band, heavily overlapping filterbank for fast-acting compression, which more closely mimics the compressive functioning of a healthy cochlea.

REFERENCES

Bachler H, Vonlanthen A (1997) Audio zoom-signal processing for improved communication in noise. Phonak Focus 18.

Boll SF (1979) Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans Acoust Speech Sig Proc 27: 113-120.

Buus S, Florentine M (1985) Gap detection in normal and impaired listeners: the effects of level and frequency. in Time Resolution in Auditory Systems (ed. Michelsen A), 159-179, Springer-Verlag, New York.

Dirks DD, Morgan D, Dubno JR (1982) A procedure for quantifying the effects of noise on speech recognition. J Speech Hear Dis 47: 114-123.

Dreschler WA (1989) Phoneme perception via hearing aids with and without compression and the role of temporal resolution. Audiol 28: 49-60.

Edwards BW, Struck CJ (1996) Device characterization techniques for digital hearing aids. J Acoust Soc Am 100: 2741.

Edwards BW, Struck CJ, Dharan P, Hou Z (1998a) New digital processor for hearing loss compensation based on the auditory system," Hear Jour 51(8): 38-49.

Edwards BW, Struck CJ, Dharan P, Hou Z (1998b) Signal processing algorithms for a new, software-based, digital hearing device. Hear Jour 51(9): 44-52.

Erber NP (1972) Speech-envelope cues as an acoustic aid to lipreading for profoundly deaf children. J Acoust Soc Am 51: 1224-1227.

Fabry DA, Van Tasell DJ (1990) Evaluation of an articulation-index based model for predicting the effects of adative frequency response hearing aids. J Speech Hear Res 33: 676-689.

Festen JM, van Dijkhuizen JN, Plomp R (1993) The efficacy of a multichannel hearing aid in which the gain is controlled by the minima in the temporal signal envelope. Scand Audiol 38: 101-110.

Fowler EP (1936) A method for the early detection of otosclerosis. Arch Otolaryngol 24: 731-741.

French-St.George M, Wood DJ, Engebretson AM (1993) Behavioral assessment of adaptive feedback cancellation in a digital hearing aid. J Rehab Res and Devel 30: 17-25.

Glasberg BR, Moore BCJ (1992) Effects of envelope fluctuations on gap detection. Hear Res 64: 81-92.

Greenberg J, Zurek P (1992) Evaluation of an adaptive beamforming method for hearing aids. J Acoust Soc Am 91: 1662-1676.

Hawkins DB, Yacullo WS (1984) Signal-to-noise ratio advantage of binaural hearing aids and directional microphones under different levels of reverberation. J Speech Hear Dis 49: 278-286.

Hoffman MW, Trine TD, Buckley KN, Van Tasell DJ (1994) Robust adaptive microhpone array processing for hearing aids: Realistic speech enhancement. J Acoust Soc Am 96: 759-770.

Houtgast T, Steeneken HJM (1985) A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. J Acoust Soc Am 77: 1069-1077.

Killion MC (1997) Hearing Aids: Past, present, future: Moving toward normal conversations in noise. Br J Audiol 31: 141-148.

Kinsler LE, Frey AR (1962) Fundamentals of Acoustics. New York: John Wiley and Sons.

Klumpp RG, Webster JC (1963) Physical measurements of equally speech-interfering navy noises. J Acoust Soc Am 35: 1328-1338.

Kochkin S (1994) MarkeTrak III identifies key factors in determining consumer satisfaction. Hear J 45(8): 39-44.

Kochkin S (1999) MareTrak V: Industry trends 1984-1997. Am Acad Audiol 11th Annual Convention, Miami, FL.

Levitt H (1991) Future directions in signal processing hearing aids. Ear Hear 12: 125-130.

Levitt H, Neuman A, Mills R, Schwander T (1986) A digital master hearing aid. J Rehabil Res Dev 23: 79-87.

Lim JS (1983) Speech Enhancement. Englewood Cliffs: Prentice Hall.

Lindemann E (1997) The Continuous Frequency Dynamic Range Compressor. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York.

Lybarger SF (1947) Development of a new hearing aid with magnetic microphone. Elect Manufact: 1-13.

Moore BCJ, Oxenham AJ (1998) Psychoacoustic consequences of compression in the peripheral auditory system. Psych Rev 105(1): 108-124.

Moore BCJ, Wojtczak M, Vickers DA (1996) Effects of loudness recruitment on the perception of amplitude modulation. J Acoust Soc Am 100: 481-489.

Nelson DA, Schroder AC (1996) Release from upward spread of masking in regions of high-frequency hearing loss. J Acoust Soc Am 100: 2266-2277.

Neuman AC, Schwander TJ (1987) The effect of filtering on the intelligibility and quality of speech in noise. J Rehab Res Dev 24: 127-134.

Oxenham, Moore BCJ, Vicker DA (1997) Short-term temporal integration: evidence for the influence of peripheral compression. J Acoust Soc Am 101: 3676-3687.

Pekkarinen E, Salmivalli A, Suonpaa J (1990) Effect of noise on word discrimination by subjects with impaired hearing, compared with those with normal hearing. Scand Audiol 19: 31-36.

Peterson PM (1989) Adaptive array processing for multiple microphone hearing aids. Ph.D. Thesis, Depart Elect Eng and Comp Sci, MIT.

Plomp R (1978) Auditory handicap of hearing impairment and the limited benefit of hearing aids. J Acoust Soc Am 63: 533-549.

Plomp R (1988) The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. J Acoust Soc Am 83: 2322-2327.

Plomp R, Mimpen AM (1979) Speech-reception threshold for sentences as a function of age and noise level. J Acoust Soc Am 66: 1333-1342.

Punch JL, Beck LB (1986) Relative effects of low-frequency amplification on syllable recognition and speech quality. Ear Hear 7: 57-62.

Rankovic CM (1999) Factors governing speech reception benefits of adaptive linear filtering for listeners with sensorineural hearing loss. J Acoust Soc Am 103(2): 1043-1058.

Roberts M, Schulein R (1997) Measurement and intelligibility optimization of directional microphones for the use in hearing aid devices. 103rd Meeting of the Aud Eng Soc, New York.

Ruggero MA (1992) Responses to sound of the basilar membrane of the mammalian cochlea. Curr Opin Neurobiol 2: 449-456.

Ruggero MA, Rich NC (1991) Furosemide alters organ of Corti mechanics: evidence for feedback of outer hair cells upon basilar membrane. J Neurosci 11: 1057-1067.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science 270: 303-304.

Soede W, Berhout A, Bilsen F (1993) Assessment of a directional microphone array for hearing-impaired listeners. J Acoust Soc Am 94: 799-808.

Stein LK, Dempesy-Hart D (1984) Listener-assessed intelligibility of a hearing aid self-adaptive noise filter. Ear Hear 5: 199-204.

Thompson SC (1999) Dual microphones or directional-plus-omni: Which is the best? in *High Performance Hearing Solutions Vol. 3; Suppl to Hear Rev* (eds. Kochkin S and Strom KE): 31-35.

Turner CW, Hurtig RA (1999) Proportional frequency compression for listeners with sensorineural hearing loss. J Acoust Soc Am, in press.

Tyler RS, Kuk FK (1989) The effects of "noise suppression" hearing aids on consonant recognition in speech-babble and low-frequency noise. Ear Hear 10: 243-249.

Tyler RS, Baker LJ, Armstrong-Bednall G (1983) Difficulties experienced by hearing-aid candidates and hearing-aid users. Br J Audiol 17: 191-201.

Uzkov AI (1946) An approach to the problem of optimum directive antenna design. C R Acad Sci USSR 35: 35.

Valente M, Fabry DA, Potts LG (1995) Recognition of speech in noise with hearing aids using dual-microphones. J Am Acad Audiol 6: 440-449.

Van Tasell DJ, Crain TR (1992) Noise reduction hearing aids: Release from masking and release from distortion. Ear Hear 13: 114-121.

Van Tasell DJ, Fabry DA, Thibodeau LM (1987a) Vowel identification and vowel masking patterns of hearing-impaired subjects. J Acoust Soc Am 81: 1586-1597.

Van Tasell DJ, Larsen SY, Fabry DA (1988) Effects of an adaptive filter hearing aid on speech recognition in noise by hearing-impaired subjects. Ear Hear 9: 15-21.

Van Tasell DJ, Soli SD, Kirby VM, Widin GP (1987b) Speech waveform envelope cues for consonant recognition. J Acoust Soc Am 82: 1152-1161.

Weiss MR, Aschkenasy E, Parsons TW (1974) Study and development of the INTEL technique for improving speech intelligibility. Nicollet Scientific Corp, Final Rep NSC-FR/4023.

Wojtczak M (1996) Perception of intensity and frequency modulation in people with normal and impaired hearing. In: Kollmeier B (ed) Psychoacoustics, Speech, and Hearing Aids. Singapore: World Scientific Publishing Co, pp. 35-38.

Wolf RP, Hohn W, Martin R, Powers TA (1999) Directional microphone hearing instruments: How and why they work. in *High Performance Hearing Solutions Vol. 3; Suppl to Hear Rev* (eds. Kochkin S and Strom KE): 14-25.

Yates GK (1995) Cochlear structure and function. In: Moore BCJ (ed) Hearing: Academic Press, San Diego.

.