Noise-reduction circuitry in hearing aids: (2) Goals and current strategies

By Donald J. Schum

Last month, the author examined noise-reduction efforts outside the hearing aid industry. In this, the conclusion of a two-part article, he discusses the two major goals of noise reduction and then describes and compares the two principal types of noise-reduction strategies.

Note: The article refers to audio samples. These samples are contained on a Noise Reduction demonstration CD, which serves as an audio accompaniment to the descriptions and figures in this article. The CD is available free of charge from Oticon, Inc, by calling 800/526-3921.

To understand how the challenge of controlling the effects of noise in hearing aids has evolved, it is important to specify the goals of noise-control circuits. The most obvious goal is to remove the noise signal without affecting the speech signal. We’ll call this Goal #1.

As discussed in our previous article, this is not easily achieved in hearing aids. That does not mean that circuitry has no effect on the performance of amplification in noisy environments. However, we must have a realistic understanding of what the effect can be. A more reasonable objective for noise-reduction circuitry in hearing aids is Goal #2, to improve the acceptability of hearing aids in noisy environments.

In other words, if users accept that the circuit cannot remove all the noise and leave only the speech, they can obtain significant benefits from the action of the processor. Noise-reduction circuitry can typically reduce the total loudness of noisier environments. This increases user acceptance of hearing aids. For, even though speech understanding is not improved, the annoyance and fatigue associated with using hearing aids may be reduced.

REducing low-frequency noise

Traditionally, it has been assumed that noisy environments are dominated by low-frequency energy. This assumption is accurate for most “environmental” noises (traffic, machinery, etc.). However, in many other situations that are difficult for persons with hearing loss, the background noise is largely unwanted speech. Despite that, previous attempts at noise control have generally focused on reducing amplification in the low frequencies. The assumptions behind these approaches are:

❖ Reduction of low-frequency amplification minimizes the effects of excessive upward spread of masking.
❖ Reduction in low-frequency amplification provides listeners with a signal that contains the frequency regions most important to their overall understanding and is more likely noise-free.

Concerning the first assumption, excessive upward spread of masking is a problem for only some listeners with sensorineural hearing loss. Also, there are only limited conditions in which reducing low-frequency gain improves speech understanding.

As to the second assumption, a reduction in gain will never improve the amount of speech that is audible to the patient. It can only reduce information. It has become clear that the dominant factor in speech understanding by a person with sensorineural hearing loss is the amount of information that is audible (i.e., above the listener's threshold and above the noise background). A broad range of frequency responses can meet that requirement for a given listener.

A decade or more ago when many hearing aids were designed to address the understanding-in-noise problem by reducing low-frequency gain, they typically used one of the following strategies:

1. (1) reducing the static low-frequency response of the hearing aid,
2. (2) giving the user a switch that changed the frequency response of the hearing aid from broadband (for quiet environments) to high-pass (for noise environments), or
3. (3) having the hearing aid automatically reduce gain in the low frequencies as the overall SPL of the listening environment increased.

Since these approaches were being implemented in single-channel, linear, analog hearing aids, the low-frequency reduction was achieved by using filtering to roll off the response. In other words, the response of the hearing aid was reduced by a certain number of dB per octave below some mid-frequency set point, which meant that reduction increased as frequency decreased.

Audio sample 5 is designed to show the effect of rolling off the low frequencies to minimize noise. The sample is speech in a background of roadside noise. In the noise-reduction condition, a 12-dB-per-octave, low-frequency roll-off is applied below 1000 Hz (down to 200 Hz). Figure 7 provides the long-term spectra of the original and processed signal. As can be heard in the sound sample, this approach makes the signal sound “thin” due to loss of much of the low-frequency energy. Also, when the noise becomes much louder than the speech for a short time, the roll-off has no effect on reducing the masking effect of this noise.

Given the basic nature of these circuits, there is no reason to expect improved speech understanding. In a large, double-blinded study, Humes et al. compared the perfor-
mance of well-fitted linear amplification with and without automatic low-frequency reduction. They found no differences in either objectively or subjectively measured performance in noise.

With the advent of multi-channel, wide dynamic range compression (WDRC) in the 1990s, low-frequency-dominated noise environments could be processed more effectively. In these multi-channel instruments, the typical reaction in a noisy environment is not to roll off the low-frequency gain but to apply greater compression in the low-frequency channel(s). This solution does not do a better job of increasing speech audibility, but it can reduce the loudness of the low-frequency dominant noise while having less of a thinning effect on the sound quality.

Audio sample 6 demonstrates the effect of two-channel WDRC compression on a speech-in-noise signal. The speech is presented against a sample of subway train noise. The signal is split at 1500 Hz. The high-frequency channel is amplified and compressed with a 2:1 compression ratio (CR). The low-frequency section is compressed using a CR of 4:1. The long-term spectra of these two signals is shown in Figure 8. The effect is to reduce the relative weighting of the noise-dominated low frequencies compared to the more information-loaded higher frequencies. Compared with audio sample 5, there is less loss of low-frequency loudness.

Nearly all advanced hearing instruments being produced today implement multi-channel WDRC processing. Thus, whether or not the hearing aid has a dedicated noise-control circuit, it will have this multi-channel response to loud, low-frequency-dominated noise environments.

This fact has two important implications. First, before the effect of the noise-control circuit is taken into account, the gain and compression parameters of the multi-channel hearing aid should be properly adjusted to ensure acceptability in noisy environments. If a noise-control circuit is needed to avoid discomfort in noise, then the core settings of the instrument are not correct.

Secondly, the patient’s impression of the response of the hearing aid to background noise may be attributable to either the effect of the noise-control circuit itself or to the core, multi-channel, WDRC action. If a patient likes how the hearing aid makes background noise seem not so loud, that may be due simply to the multi-channel WDRC effect.

**DSP-BASED APPROACHES**

The introduction of DSP-based instruments in the mid-1990s led to advances in areas such as feedback cancellation and directionality. However, when it comes to noise control, hearing aid manufacturers can only implement those algorithms that exist. The fact that hearing aids are now on a DSP platform does not mean that the hearing industry has a solution to the noise problem.

As noted last month, other much larger and better-funded industries have searched for an effective algorithm to remove noise from a speech signal—Goal #1. They have had only limited success. On the other hand, when it comes to Goal #2, improving acceptability in noisy environments, the DSP platform has led to significant progress.

**Modulation detection**

Several of the DSP-based instruments on the market use some variation of modulation detection to classify the input as speech or noise. In this scheme, the ongoing amplitude modulations of the input signal are monitored. Speech in quiet is known to have relatively deep (15 dB or greater) modulations at a rate between...
approximately 3 and 10 Hz. This modulation pattern reflects the syllabic structure of speech: three to six syllables per second, each typified by a vowel (intense, harmonically structured) with one or more consonants (often less intense and aperiodic). In contrast, environmental sounds tend to be more stable in terms of ongoing amplitude. It is unusual for a non-speech noise source to have a modulation rate and depth similar to that of speech.

As implemented in hearing aids, the input signal is divided into multiple channels and the modulation behavior is monitored in each channel. If the modulation rate and depth are like those of speech in quiet, then that channel is passed without gain reduction. If the modulation behavior in the channel is more stable, it is assumed that the channel is dominated by steady-state noise and gain reduction is applied.

It is important to note that there is a difference between being able to determine if a given channel is dominated by speech or by noise (classification) and being able to separate speech from noise. Even if the modulation detection circuitry is accurate in finding a noise-dominated channel, the algorithm is unable to separate the noise from the speech in that channel. Since speech is broadband, any channel identified as being dominated by noise will still have speech information, but at a relatively poor S/N.

As implemented in hearing aids, modulation-detection algorithms are a version of spectral subtraction. The modulation detector determines the approximate frequency content of the noise (as precisely as the width of the channels used). When channels with a strong noise component are identified, they are removed (to some degree) via gain reductions from the broadband response. As in all spectral subtraction approaches, the effect on the speech signal is directly proportional to the frequency content of the noise: the more broadband the noise, the greater the effect on speech.

Audio sample 7 shows the effect of a modulation-detection approach to noise control for a narrowband competition. Speech is presented against the hissing sound of a milk steamer in a coffee boutique. The hiss is centered about 3200 Hz (see Figure 9, blue line). A third octave band filter centered at 3200 Hz is used to reduce that band by 30 dB (Figure 9, red line). The auditory effect is a reduction in the annoyance of the milk steamer.

It is clear from audio sample 7 that modulation detection can be effective at meeting Goal #2, improving acceptability in the presence of band-limited competition. However, many of the most difficult listening situations that commonly face persons with hearing loss have broadband competition.

Audio sample 8 provides six different real-world background sounds. Figure 10 provides the long-term spectra for these environments. The purpose is to demonstrate that most common competition tends to be broadband. Given that the long-term spectrum of speech falls off by about 6 dB/octave above 500 Hz, the only one of these backgrounds that would not overlap speech in the important mid- and high frequencies is the “subway” environment. The broadband nature of most common communicative environments significantly limits the ability of modulation-detection approaches to minimize
the loudness of noise without significantly affecting the speech signal.

Audio sample 9 shows the effect of modulation detection on speech presented in a background of street-side noise. This noise sample had a broad bandwidth (Figure 11, blue line) similar to the spectrum of the speech signal (green line). When the speech and noise were mixed, a filter that extends through the frequency range with minimal modulations (3000 Hz) was applied and this part of the spectrum was reduced by 20 dB. The spectrum of the filtered speech-plus-noise signal is presented as the red line in Figure 11. The auditory effect is obvious: The loudness of the noise is reduced, but so is a significant amount of the speech energy.

The effectiveness of modulation detection in classifying signals is limited since this approach cannot distinguish speech in noise at a poor S/N from pure noise. Speech in quiet is characterized by differences of 15 dB, 20 dB, or more between the most intense phonemes and the least intense segments. However, when speech is in a stable background noise at a S/N of, say, +5 to +10 dB, the deep modulation characterized by speech in quiet is lost, filled in by the stable background noise. Although there is still plenty of speech information above the background noise level and therefore available to the listener, a classification scheme that looks at modulation rate and depth would identify this segment as noise and reduce gain, thereby sacrificing audibility.

The strongest modulations in speech are in the lower frequencies. Figure 12 provides the waveforms of the combined speech-plus-noise (café) signal once filtered into octave bands (250-500 Hz, 500-1000 Hz, 1000-2000 Hz, 2000-4000 Hz, and 4000-8000 Hz). As can be seen, the depth of modulation decreases in the higher frequencies. In other words, the likelihood of a “noise” decision being made increases in the higher frequencies. Gain reductions would then be applied in the frequency regions that carry the most speech information.

Two recent studies examined the effect on speech understanding in noise of actual, advanced DSP-based products that implement modulation detection.10,11 In other words, can modulation-detection-based systems meet Goal #1 (remove noise, leave speech). In both studies, when patient performance was compared with and without noise control activated, there were no differences in objectively measured speech understanding.

Clinicians who recognize that modulation-detection systems cannot meet Goal 1 will sometimes use the technology anyway to achieve Goal 2 (improve acceptability). They report patients saying things such as “When I enter a noisy room, the hearing aid gets quieter.” Experimentally, it has been shown in some cases that these systems can provide improved subjective ratings of aided comfort or acceptability.12 (This subjective finding has not been found in all studies.) However, patient response to hearing aid performance deserves close examination.
It must be remembered that modulation-detection systems have almost always been implemented in advanced multi-channel, non-linear hearing aids. Even without their noise-reduction system activated, such fully automatic hearing aids should not be uncomfortable and should reduce gain in a louder environment.

Remember also that, since the ability of a noise-reduction algorithm to identify and classify noise will never be perfect, it should not be used as a mechanism to protect against loudness discomfort. It is one thing to use such a system to reduce loudness of noise if the patient prefers not to have to listen when the speech information is harder to detect. It is quite another to have to operate the system to make sure the signal does not become uncomfortable. Further, if the patient notices that the gain of the device drops when the environment gets louder, he or she may simply be hearing the fully automatic non-linear action. Using modulation detection to reduce the loudness of noisy environments is a legitimate strategy. However, both the clinician and the patient should recognize the true nature of this type of circuitry.

**Synchrony detection**

A newer classification approach is synchrony detection. It is based on the fact that, while speech energy is distributed across the full bandwidth of the hearing aid, the pattern of energy in different frequency regions is precisely timed with the periodic action of the vocal folds (fundamental frequency). Since voiced speech (vowels and voiced consonants) is composed of a series of harmonics of the fundamental frequency, the amplitude in the mid- and high frequencies is driven by the amplitude of the periodic vocal fold vibration. Every time the vocal folds open and close (100 to 250 times a second, depending on the pitch of the speaker’s voice), a broadband pulse of energy is created. Figure 13 shows the on-going amplitude fluctuations in the very low frequencies (reflecting the fundamental frequency) and in four high-frequency bands, each 1000 Hz wide.

A synchrony detector searches for this precisely timed, synchronous pattern of energy in the higher frequencies by tracking the on-going correlations of instantaneous amplitude across frequency regions. This robust speech-detection system has been demonstrated to be sensitive to the presence of speech in broadband competition all the way down to 0 dB S/N and beyond. Commercially, this approach has been used only in the VoiceFinder circuitry of Adapto by Oticon.

Just as with modulation detection, one must distinguish between the ability of synchrony detection to classify speech versus noise and its ability to separate the two. Synchrony detection is no better at separating speech from background noise than any other noise circuit now available. However, in pursuit of the second goal of noise control, VoiceFinder with synchrony detection uses the classification in a unique manner. Whenever speech is detected, the system provides full amplification and compression characteristics in accordance with the non-linear fitting rationales implemented in the Adapto product (see Schum and Pogash for a description). This full amplification state is referred to as the speech mode.

When synchrony analysis determines that no speech is present, the device transitions into the comfort mode and applies greater compression to reduce the loudness of non-speech background noise. It is assumed that this action reduces the long-term fatigue of using amplification. Audio sample 10 shows the effect of VoiceFinder. The speech signal is presented against a background of traffic noise. Initially, the level of the noise is reduced as in the comfort mode. Once the speech starts, the hearing aid transitions into the speech mode and the gain is immediately increased (including the gain applied to the background noise). A few seconds after the speech stops, the level of the noise begins to decrease gradually as the hearing aid transitions back into the comfort mode.

**COMPARING APPROACHES**

It is of interest to evaluate the practical differences between the two existing alternative approaches. There are two dimensions that the clinician should consider when evaluating these systems.

The first of these is the relative sensitivity to the presence of speech versus the presence of noise. As indicated earlier, once S/Ns become poorer than about +10 to +15 dB, modulation-detection systems will classify a mixed speech-plus-noise signal as noise and then drop the gain in that channel. Patients may prefer that if they do not care to search the mixed signal for...
whatever speech information may be available. However, many patients may still be able and may still wish to extract meaningful information at such S/Ns. The gain drops could seriously impair the audibility of speech information.

Synchrony detection is more resistant to the presence of background noise. The upper panel of Figure 14 provides the response of a synchrony-detection system (VoiceFinder) in various backgrounds. These measures are based on the real-ear response of the circuit (as measured on an acoustic manikin). The metric is the reduction in gain when the noise-processing circuit is activated. For each environment, the left-hand bar is for 500 Hz (labeled “LF”) and the right bar is for 2000 Hz (labeled “HF”). Nine environments were tested. The seven on the right contain speech or speech-like signals (music). The two conditions on the right did not contain speech.

Based on the design, there should be no changes in level if speech is detected and the gain should drop if speech is not detected. As can be seen, the drops in gain occur only for the two conditions without speech: traffic noise and unmodulated speech-shaped noise. For the other conditions, including speech in noise down to 0 or –2 dB S/N, no gain drops were recorded, indicating that the synchrony detector could still identify the presence of a speech signal.

For comparison purposes, the middle and bottom panels of Figure 14 show the same analysis for two models of hearing aids that incorporate modulation detection. In contrast with the upper panel, gain drops occur in both the low frequencies and (especially) the high frequencies in nearly all of the environments. In general, whenever the input level is in the moderate-to-high range and there is any noise present, these systems apply gain reductions. As shown above, modulation-detection systems tend to be attuned to the presence of noise whereas synchrony detection tends to be attuned to the presence of speech.

Secondly, the clinician needs to decide what sort of response he/she wants the system to have in a noisy background. If a mixed speech-plus-noise signal is present, what should the hearing aid do? Assuming that the core, multi-channel, non-linear action of the device would amplify this mixed signal to an acceptably comfortable level, do you want further gain reductions?

Most non-linear fitting rationales are designed to present aided speech at levels that provide good audibility and acceptable comfort, even in the case of the higher input levels typical of noisy environments. The question is whether or not, when speech is present at a moderate S/N such as +5 or +10 dB, one wants to give the patient further gain reductions.

If the noise is of a relatively limited bandwidth, using modulation detection with the gain reductions and the associated drop in the loudness of the noise would not have a major impact on the remaining speech and would probably be sensible. However, if the noise is broader in bandwidth, as is typical of many everyday listening situations, then the gain reductions used in modulation-detection noise control would detract from available speech information. In contrast, a synchrony-detection system would not generate gain changes until the speech signal was no longer detected.

How do patients want their hearing aids to perform? The answer probably depends on many factors. One of these is how important it is to the patient to understand the speech. If the patient is not highly motivated to extract speech information from a challenging listening environment, then the gain reductions characteristic of modulation-detection systems are probably acceptable.

However, in a recent study, Souza and Kitch asked listeners to set their preferred gain level in a hearing aid when listening to both speech in quiet and a mixed speech-plus-noise signal (+8 dB S/N). The listeners were instructed to pick the level that provided the best intelligibility. The signal was presented via three types of amplification circuits: linear, output-limiting compression, and WDRC. The comparison of preferred gain level for speech in quiet versus speech in noise is presented in Figure 15. As can be seen, regardless of the type of amplification or input level, on average the subjects pre-
ferred the same or even slightly higher gain levels when listening in noise than when listening in quiet. When the task was to hear and understand speech, these subjects did not want gain reductions. Rather they preferred full bandwidth audibility of the signal.

CONCLUSION
Patients with sensorineural hearing loss need significant assistance in noisier situations. Technologies exist (directional microphones and FM systems) that can provide a more advantageous signal to the patient. Furthermore, clinicians can tell patients and their families about helpful behavioral strategies.

Despite these efforts, a great need remains for circuitry that can improve speech understanding in noise. Significant resources are being applied to this challenge both within the hearing aid industry and beyond. Will progress take place? Almost certainly. Will we see dramatic improvements in the near future? Probably not.

More likely, we will see incremental improvements year by year. DSP designers will find more creative ways to analyze and treat these complex signals. In the future, the currently sophisticated techniques of modulation and synchrony detection will seem simplistic.

In closing, improving speech understanding in noise is both a high priority and a difficult challenge. Serious, concerted efforts will continue as long as it takes to overcome that challenge and to meet that priority.

Donald J. Schum, PhD, CCC-A, is Vice-President, Audiology and Professional Relations, Oticon, Inc. Correspondence to Dr. Schum at DJSchum@Oticonus.com.

REFERENCES