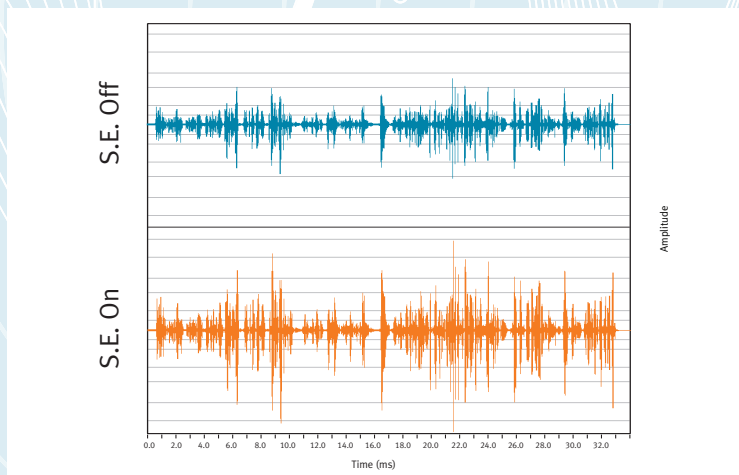




## Speech Enhancement: SURPRISING BENEFITS FOR SOFT AND DISTANT SPEECH



### Executive Summary

Spectral enhancement algorithms have been shown to provide some limited benefit for speech perception<sup>1-6</sup>. However, their benefit has sometimes been limited by sound quality issues related to loudness and distortion<sup>1, 7, 8</sup>. With the introduction of Element™ 16 and Element™ 8, Unitron Hearing has demonstrated a new approach to speech enhancement, as well as new ways to use it. Distortion is eliminated by controlling the application of speech enhancement at high output levels. Speech enhancement can now be used to hit fitting targets for speech, while providing less than target gain for noise by altering the gain model at the fitting. This approach has been taken to its logical conclusion using level dependent speech enhancement in Element 16. Speech enhancement LD (level dependent) allows the wearer to obtain significantly more gain for soft speech, and less gain for noise than was ever possible before.

## Early Speech Enhancement Algorithms

Enhancing the amplitude for speech has been attempted in several different ways<sup>1-5, 7, 9-11</sup>. If it is possible to recognize when speech is present, and provide more gain for that speech than the surrounding environmental sounds, both clarity and comfort are improved. However, past attempts in the laboratory have been shown to improve either intelligibility or sound quality, but rarely both<sup>4, 7, 8</sup>. Those issues arose before implementing the algorithm in a wearable hearing aid. Further limitations to the dynamic range are imposed by hearing aid components operating on battery power. The restrictions imposed by operating the algorithm on a hearing aid in real-time imposes limitations to the dynamic range and processing power, as shown by the example in Figure 1a.

The left panel, Figure 1a, shows a single frequency I/O function for a typical hearing instrument (black line). The blue area of the graph demonstrates the gain increase that would occur for speech with an early speech enhancement algorithm. Since the enhancement is not level dependent, the same gain increase is applied to speech regardless of the input level. If the hearing aid is fitted using a traditional gain model such as NAL-NL1 or DSL[i/o], the enhanced output will be above target. Exceeding target can be beneficial for soft speech, and it may also be acceptable for

average speech levels. However, it will be very uncomfortable for loud speech. Furthermore, loud speech will saturate the hearing aid at a lower input level when speech enhancement is engaged. Notice that when speech enhancement is applied, it changes the input level at which saturation is reached. The saturation point has moved to the left in Figure 1a when speech enhancement is on. This saturation and the over-amplification of loud speech can contribute to sound quality issues for speech enhancement. Some minor changes to the performance of the enhancement algorithm and the gain model to which it is applied can overcome these limitations.

## Speech Enhancement in Element™ 8:

Early enhancement algorithms were limited because they increased the likelihood of saturation distortion and provided too much gain for loud speech as in Figure 1a. Element 8 does not increase the likelihood of saturation distortion because gain is rapidly reduced as the output of the hearing instrument approaches saturation. The gain function of Element 8 is shown in Figure 1b. Like the previous model in Figure 1a, enhancement is applied equally over a broad range of input levels. However, unlike the model in Figure 1a, the input at which saturation occurs for speech does not move to the left because of the rapid diminution of gain as the output of the hearing aid nears

Figure 1a

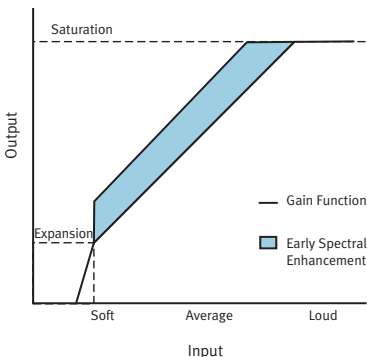


Figure 1b

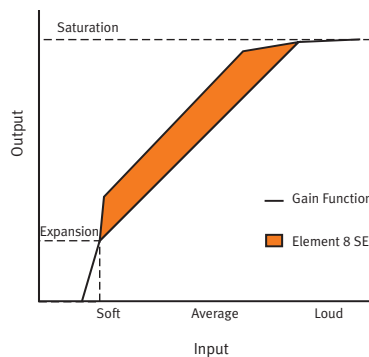
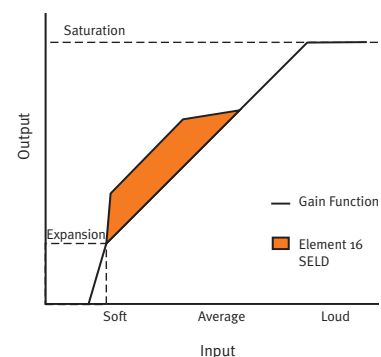


Figure 1c



saturation. Thus, the speech enhancement in Element 8 does not increase saturation distortion for the wearer.

The issue of loudness tolerance for higher level speech inputs is handled by applying an updated view of the goals of target gain rules. For example, target gain rules such as DSL[i/o] and NAL-NL1 calculate optimal hearing aid gain. However, optimal hearing aid gain can be defined with more than one goal in mind. Possibly it means the level at which speech perception has been optimized, or maybe it refers to the point at which normal loudness perception has been restored. Regardless of the goal, each prescriptive formula requires gain calculations at several frequencies to achieve desired gain/output targets. But have we ever asked the question, gain for what? Gain was simply gain. The addition of 6 dB for speech meant the addition of 6 dB for everything else. However, a speech enhancement algorithm that correctly applies a gain boost for speech, and not for the surrounding noise, requires a more sophisticated definition of gain.

By turning speech enhancement off or on, it is now possible to apply the same amount of gain for all sounds including speech or to apply up to 6 dB of additional gain just for speech. Therefore, the way in which a gain target is applied depends on the goal of the wearer. In Figure 1, the i/o function of the hearing instrument was a match to target and any speech enhancement that was applied provided gain in excess of the target level, but that is not a requirement. Applying less than the calculated target gain and then providing enough enhancement to reach the target in the presence of speech is also possible. The fitter must then choose how he/she would like to reach the correct amount of target gain. The options include:

- **Traditional Fitting** – Quick fit the aid to fully match target without any speech enhancement. This provides the same gain for speech and environmental sounds.

- **Speech Preference** – Quick fit the aid a few dB below a full match to target, then apply enough speech enhancement to match target for speech only. Preferential treatment is given to speech, keeping the level of environmental sounds down.

- **Speech Boost** – Quick fit the aid to fully match target, then add additional gain for speech using enhancement. This can considerably boost speech for wearers who like a lot of power, while maintaining lower gain levels for environmental sounds.

This fitting approach for speech enhancement opens up new possibilities to fitters which were never previously available. The effect can be seen in U:fit™'s new quick fit logic for Element 16, Element 8 and Element 4. Element 16 and Element 8 are quick fit to a lower setting on the adaptation manager than Element 4. The adaptation manager reduces the overall gain of the hearing instrument during the quick fit relative to the full target gain rule for an experienced wearer. The initial gain reduction for Element 16 and Element 8 is offset by engaging speech enhancement. The net effect is a quick fit directly to a “speech preference” type fitting. Normal loudness and clarity for speech are maintained since the gain model provides amplification equal to the full target gain rule application. However, non-speech signals are significantly de-emphasized eliminating complaints of overly amplified background noise. Since Element 4 has no speech enhancement algorithm, more initial gain is applied during the quick fit.

For cases where the wearer desires more gain, the gain boost approach allows the fitter to quick fit to full target, and then apply speech enhancement. While most wearers might find this a bit too loud, those who are long time hearing aid wearers or those with severe losses may find this approach ideal.

In summary, rolling off the enhancement applied to loud speech that approaches saturation eliminates saturation

distortion. The gain increase for speech due to speech enhancement is accounted for by more sophisticated quick fit logic. Thus, more appropriate amplification is provided for speech while eliminating loudness complaints due to over amplification. In addition, speech enhancement can be used to boost speech for power users, slightly in excess of the full target gain rule.

## Speech Enhancement with Level Dependence (LD) – Element™ 16:

The two principles described above allow Element 8 to provide the benefits of speech enhancement without any of the associated loudness or distortion problems. Rolling off the speech enhancement applied to near saturation speech outputs is one small step toward level dependent enhancement. However, if the point at which the speech enhancement rolls off is lowered to the level of average conversational speech, a more powerful application becomes possible (see Figure 1c). By applying speech enhancement only to soft speech, considerable gain can be added for speech while significantly limiting the audibility of surrounding soft noises. Even more gain can be provided by the level dependent speech enhancement in Element 16 than that in Element 8 because louder speech gets no enhancement at all.

For a hearing aid to apply level dependent speech enhancement, its performance must satisfy the following three criteria. The hearing instrument must:

- 1) correctly apply additional gain to bands dominated by speech, but not bands dominated by noise;
- 2) accurately detect speech in quiet and in noise; and
- 3) provide proportionately more gain for speech at some desired input levels than for speech at other less desirable input levels.

Once these three conditions are met, the hearing aid is providing truly level dependent and adaptive speech enhancement. Here are some examples of how speech enhancement LD works in Element 16.

## Correct Detection and Application

The first step in the application of any adaptive feature in a digital hearing aid is correct detection. For speech enhancement, correct detection requires the ability to recognize speech when it is present. By extension, this infers the ability to recognize when speech is not present. For a speech enhancement algorithm, both requirements are equally important. Speech enhancement should only be applied when speech is present, and it should not be applied when no speech is present.

The following three figures (Figures 2, 3, and 4) will demonstrate this principle for very soft inputs. Soft inputs are used because of the level dependency of the speech enhancement in Element 16. If the same speech signal is presented to the hearing aid at different input levels, more gain is added for soft speech than average or loud speech. Hence, an example using soft speech provides the clearest evidence for appropriate detection.

Figure 2 demonstrates the impact that speech enhancement has on soft speech. The bottom panel shows the waveforms for approximately 30 seconds of an adult female voice with a long term average input level to the hearing aid of 53 dB SPL. The bottom waveform is the output of the hearing aid in a 2 cc coupler with speech enhancement turned on. The waveform just above it shows the output of the hearing aid with speech enhancement turned off. Two things are quickly apparent: the amplitude of the enhanced speech is substantially larger than the non-enhanced speech, and the noise floor (between phonemes) is the same in both waveforms.

The top panel is an FFT of the full thirty seconds from the same speech signal. The blue area shows the average amplitude of the soft speech without speech enhancement. The orange area represents the enhanced speech. In this example, there is an average increase of approximately 5 dB when speech enhancement is turned on. The momentary enhancement varies as the input level of the speech signal rises and falls. In this case, it varies around 5 dB.

Conversely, the impact of speech enhancement in the presence of noise is negligible. Figure 3 shows a recording of a 50 dB SPL traffic noise passed through the same hearing aid as that for Figure 2, and under exactly the same conditions. Once again, the bottom waveform is the output of the hearing aid with speech enhancement on. The one

above is the output with speech enhancement off. The blue FFT in the top panel shows the average amplitude of the traffic noise without speech enhancement, and the orange area shows the average amplitude with speech enhancement.

The most salient feature of Figure 3 is the lack of any discernable effect caused by speech enhancement. The average amplitude difference between the two recordings is < 0.2 dB. This is substantially different from the 5 dB gain increase that speech enhancement provided for speech in Figure 2.

In Figure 4, the speech (53 dB SPL) and traffic noise (50 dB SPL) have been presented to the hearing aid simultaneously. Therefore, a 3 dB SNR is present at the

Figure 2  
FFT

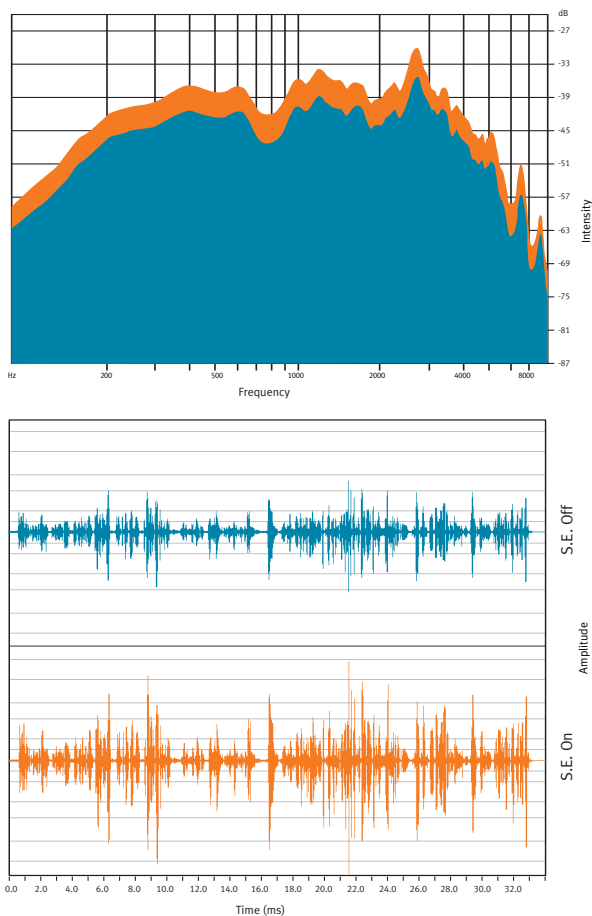
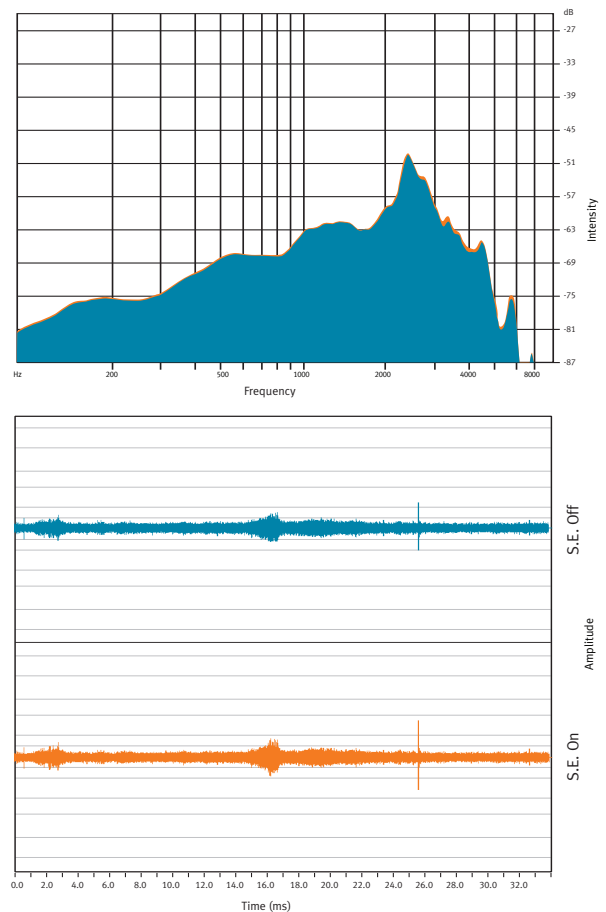


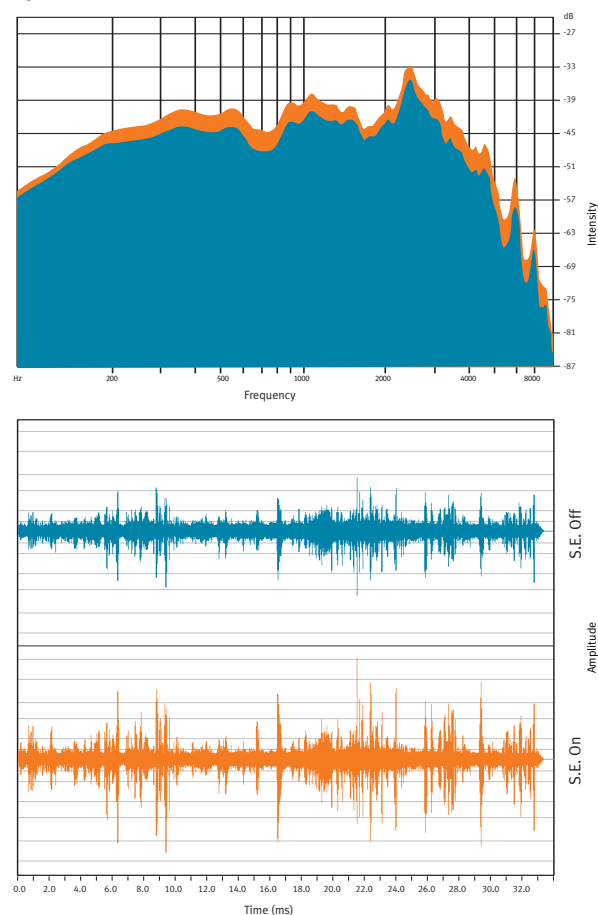
Figure 3  
FFT



hearing aid microphone, a difficult but not unrealistic SNR for a hearing impaired listener.

The layout of Figure 4 is identical to Figures 2 and 3. This time the effect of speech enhancement is obvious for any phonemes with amplitudes greater than the noise floor. When Figures 2 and 4 are compared, most of the speech peaks are nearly identical for both waveforms. This demonstrates that the algorithm correctly identifies and amplifies speech in quiet and in noise. However, comparing the noise levels of the two waveforms of Figure 4 (between the speech peaks), there is virtually no difference attributable to speech enhancement. This demonstrates that speech enhancement does not mistake noise for speech, and provides virtually no amplification for the noises that occur between phonemes.

Figure 4



The FFT's in the top panel of Figure 4 reveal the same effect. The enhanced speech has an average of 3 dB more gain than the non-enhanced speech. The slight reduction in the effect of enhancement, from 5 dB in Figure 2 to 3 dB in Figure 4, is due to the presence of the high noise floor, +3 dB SNR. Some of the speech is momentarily below the noise floor in some bands at different points in time; therefore, no enhancement is applied to those bands at those times.

## Additional Benefits of Speech Enhancement LD

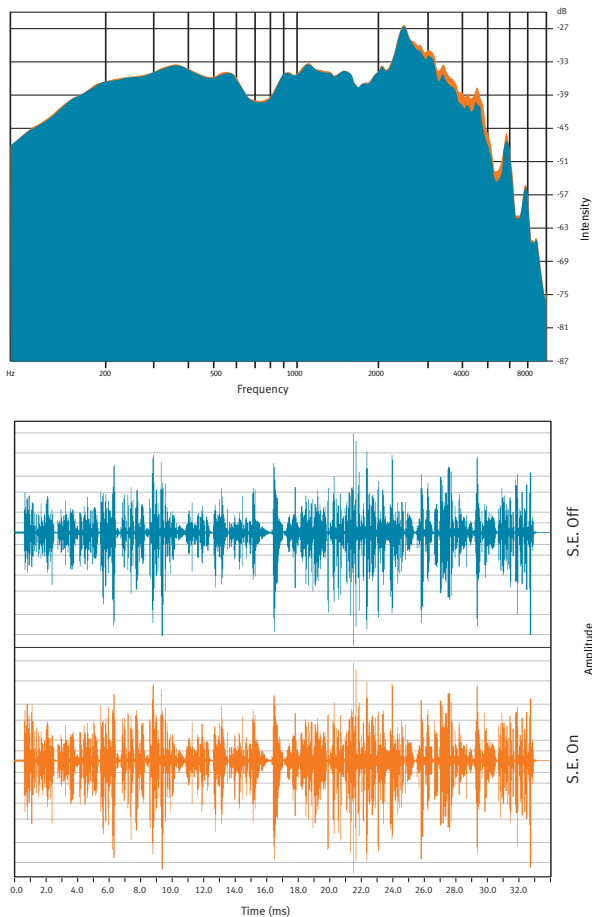
Figure 1c contains an example of speech enhancement LD. The additional gain applied to speech has a gradual onset and offset relative to the input level. Furthermore, it provides a gain boost only for soft speech; thus, the problems of loudness and saturation are avoided completely. Meanwhile, a substantial boost for soft speech is provided without concurrently over amplifying soft environmental sounds. The level dependency allows even more speech enhancement to be applied than would otherwise be tolerated by the wearer. For example, once the gain of the hearing instrument is set, adding an additional 3 – 4 dB of linear amplification for loud speech may not be tolerated by some hearing aid wearers suffering from a reduced dynamic range. However, if the additional gain were only applied to soft or distant speech, as much as 5 – 6 dB may even be welcomed. In other words, there are substantial, sometimes surprising, benefits to speech enhancement LD.

Compare Figure 2 to Figure 5 for an example of Element 16's level dependent speech enhancement. Figure 2 and Figure 5 show the effects of enhancement for the same speech passage. However, the speech in Figure 2 was presented at 53 dB SPL and that for Figure 5 was presented at 70 dB SPL. This results in a 5 dB increase in speech amplitude for

the soft speech (Figure 2) and less than a 1 dB increase for the louder speech (Figure 5). Raising the average input level of the speech any further would yield equal output amplitudes whether speech enhancement were on or off. At such high input levels, no speech enhancement is applied at all.

The same approach to target gain fittings that was discussed for Element 8 above is applied to Element 16. In this case, it is possible to provide substantial gain for soft speech. Therefore, it is even more important to apply reduced overall gain for environmental sounds at the time of the fitting. Our internal field trials have demonstrated that this new speech enhancement is so effective that reducing the overall gain at the quick fit is nearly always

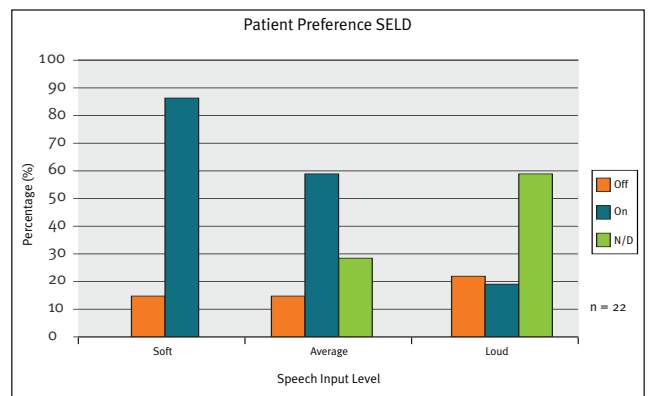
Figure 5



required, even for experienced hearing aid wearers. Figure 6 shows the results of paired comparison tests that were obtained during the internal field trials for Element 16.

When asked to provide their preference for level dependent speech enhancement (On/Off/No Difference) while listening to soft, average or loud speech signals, the results were quite compelling. Out of the 22 participants tested, nearly 90% preferred speech enhancement to be on when listening to soft or distant speech. Surprisingly, nearly 60% still preferred speech enhancement for average level speech whereas 60% of participants reported that there was no difference with speech enhancement on or off for loud speech. This is precisely the purpose of the level dependent implementation. Hearing aid wearers have lost sensitivity for soft sounds. A substantial boost for soft or distant speech can improve audibility and comfort. However, loud speech is typically well above their threshold of sensitivity and does not require an additional boost. Given that most participants could not tell the difference with the speech enhancement on or off during loud speech, this goal has been achieved. The initial surprise was the strong preference for speech enhancement when listening to average level speech. This is probably because the amplitude of the speech enhancement has already begun to drop by that level of input, and a little extra gain is deemed as acceptable and welcome under the circumstances.

Figure 6



## Summary

Spectral enhancement algorithms have historically been shown to provide some limited benefit for understanding speech in noise at specific SNRs. However, their benefit has always been limited by sound quality issues related to loudness and distortion. With the introduction of Element 16 and Element 8, Unitron Hearing has demonstrated a new approach to speech enhancement. By controlling the application of speech enhancement at high output levels, distortion has been eliminated. By altering the gain model at the fitting, speech enhancement can now be used to hit fitting targets for speech, while providing less than target gain for noise. Furthermore, speech enhancement LD in Element 16 allows the wearer to obtain significantly more gain for soft speech, and less gain for noise than was ever possible before.

## Bibliography

1. Baer, T. and B.C.J. Moore. (1993) Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality and response times. *Journal of Rehabilitation Research and Development*. 30(1): 49-73.
2. Bunnell, H.T. (1970) On enhancement of spectral contrast in speech for hearing-impaired listeners. *The Journal of the Acoustical Society of America*. 88(6): 2546-56.
3. Clarkson, P.M. and S.F. Bahgat. (1991) Envelope expansion methods for speech enhancement. *The Journal of the Acoustical Society of America*. 89(3): 1378-82.
4. Franck, B.A., et al. (1999) Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. *The Journal of the Acoustical Society of America*. 106(3 Pt 1): 1452-64.
5. Jamieson, D.G., R.L. Brennan, and L.E. Cornelisse. (1995) Evaluation of a speech enhancement strategy with normal-hearing and hearing-impaired listeners. *Ear and Hearing*. 16(3): 274-86.
6. Kollmeier, B. and R. Koch. (1994) Speech enhancement based on physiological and psychoacoustical models of modulation perception and binaural interaction. *The Journal of the Acoustical Society of America*. 95(3): 1593-602.
7. Lyzenga, J., J.M. Festen, and T. Houtgast. (2002) A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity. *The Journal of the Acoustical Society of America*. 112(3 Pt 1): 1145-57.
8. van Buuren, R.A., J.M. Festen, and T. Houtgast. (1999) Compression and expansion of the temporal envelope: evaluation of speech intelligibility and sound quality. *The Journal of the Acoustical Society of America*. 105(5): 2903-13.
9. Alcantara, J.I., et al. (1994) Preliminary evaluation of a formant enhancement algorithm on the perception of speech in noise for normally hearing listeners. *Audiology*. 33(1): p. 15-27.
10. Kates, J.M. (1994) Speech enhancement based on a sinusoidal model. *Journal of Speech and Hearing Research*. 37(2): 449-64.
11. Warren, R.M., et al. (1997) Spectral restoration of speech: intelligibility is increased by inserting noise in spectral gaps. *Perception and Psychophysics*. 59(2): 275-83.

## Author

Don Hayes, PhD, Director of Audiology

