HiResolution™ Sound Processing

Jill B. Firszt, Ph.D.
Medical College of Wisconsin, Milwaukee, WI

HiResolution™ (HiRes™) is the sound processing system implemented in the HiResolution™ Bionic Ear and programmed using the SoundWave™ software. HiRes Sound can be implemented on the CII behind-the-ear (BTE) and HiRes Auria™ ear-level processors, as well as on the bodyworn Platinum Sound Processor (PSP). HiRes is a fully integrated component of the HiResolution Bionic Ear and is made possible by recent advances in signal-processing capabilities.

What is cochlear-implant sound processing?

Sound processing defines the way a cochlear implant system transforms acoustic sound into electrical stimulation of the auditory nerve. All sound processing strategies are a combination of how the acoustic input from the microphone is analyzed for translation into electrical current, how and when current is delivered to the electrode contacts, and the stimulus waveform at each electrode contact. The eventual goal of a sound processing algorithm is to deliver a high-fidelity electrical representation of the incoming sound so that the implant user perceives the acoustic signal in a similar way to that of a normal-hearing listener, even though he or she has a severe-to-profound sensorineural hearing loss.

An accurate electrical representation of the acoustic signal is important for understanding speech, especially in challenging listening environments. Speech is a complex signal that contains frequency, intensity, and temporal cues, all of which are encoded with varying degrees of accuracy by implant systems. There is a tremendous level of redundancy in those cues so that, in a quiet environment, if some cues are missing, speech can still be understood. However, in a challenging environment, the accuracy of sound representation becomes more important because many of the cues may be masked by noise.

As an example, Figure 1 shows the spectra and time waveforms of the syllables /ba/ and /bi/. These two signals have very similar overall time envelopes, but their spectra are different. In contrast, Figure 2 shows the spectra and time waveforms of the syllables /ba/ and /wa/. For this pair of signals, the spectra are similar but the time envelopes are different. In order for an implant user to discriminate the signals in Figure 1, the implant must accurately represent the frequency information. On the other hand, for the implant user to discriminate the signals in Figure 2, the implant must accurately represent the temporal cues. Therefore, an implant system must represent signals with a high degree of accuracy in both the time and frequency domains.

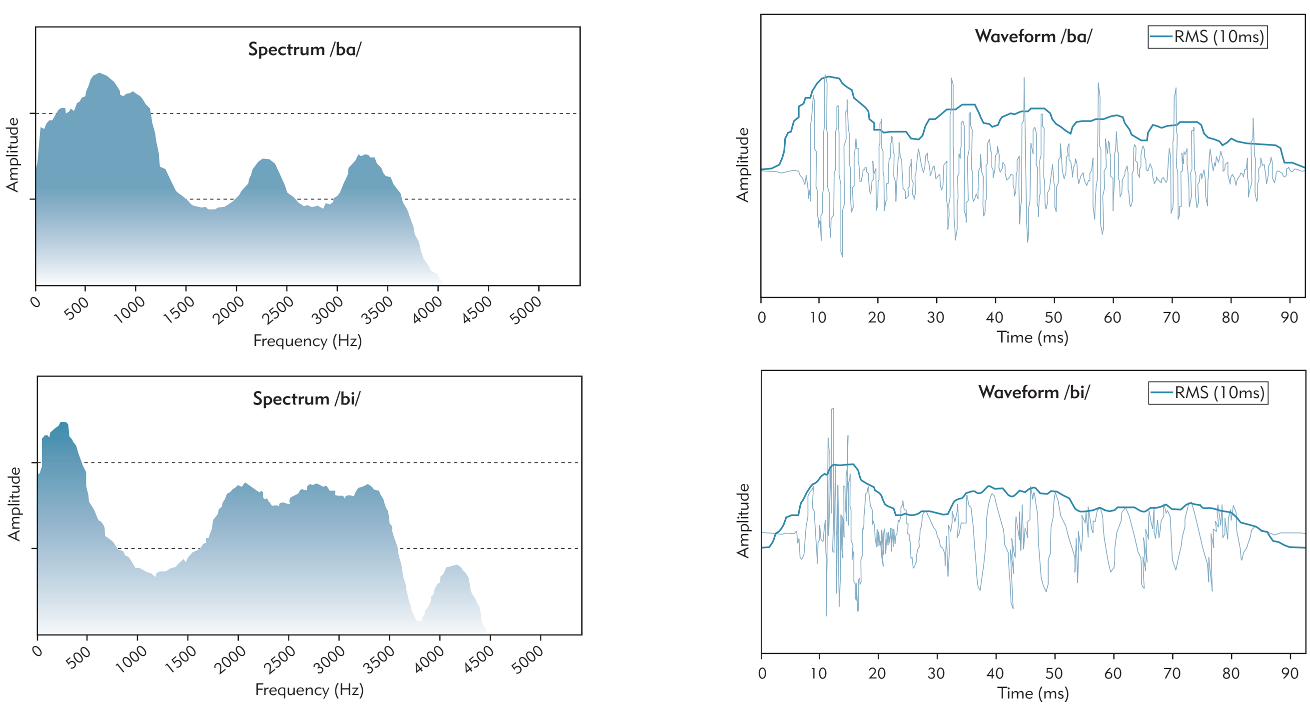

Figure 1. Spectrograms and time waveforms of the syllables /ba/ and /bi/.

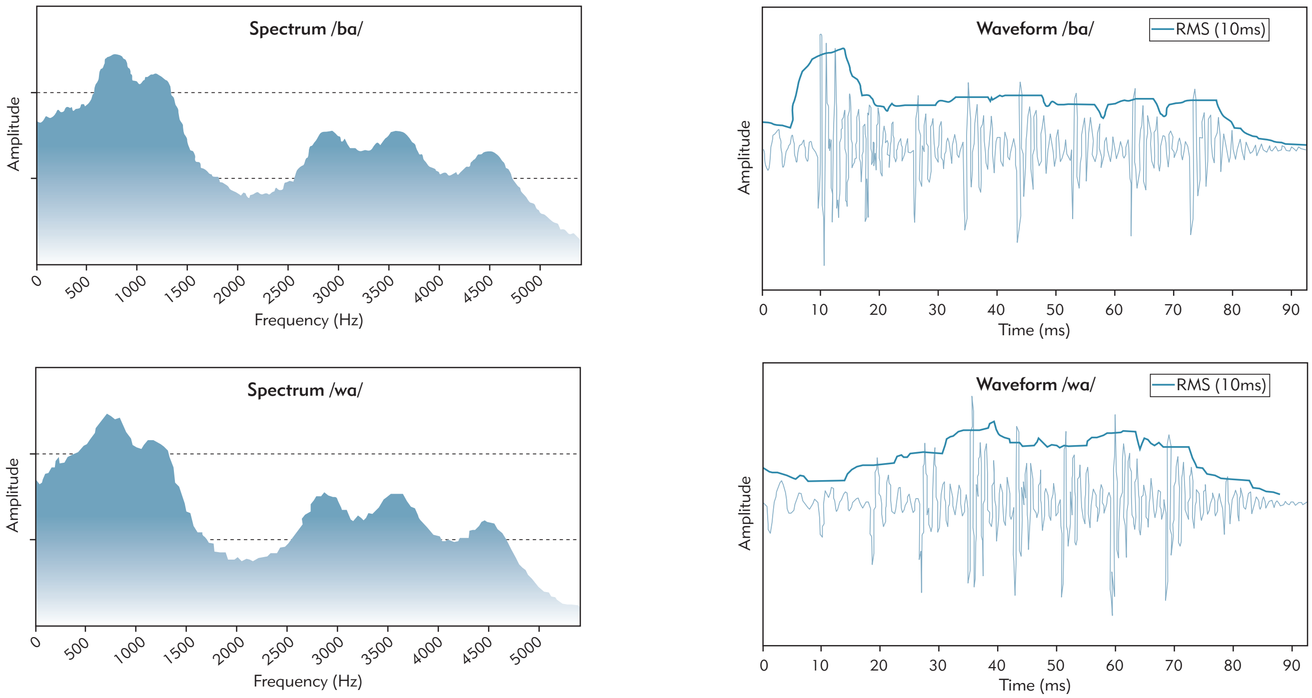

Figure 2. Spectrograms and time waveforms of the syllables /ba/ and /wa/.

HiRes sound processing was designed to represent the incoming acoustic signal with a high level of fidelity in the electrical signal delivered to the cochlear-implant electrode, and to more accurately represent spectral and temporal cues. That high-resolution signal, in turn, should facilitate better speech understanding, especially in noise, and should provide improved representation of music.

What is HiRes sound processing?

In order to understand HiRes sound processing, it is helpful to go through the steps of what happens between when sound is captured by the implant system and when electrical stimulation is delivered to the electrode implanted in the cochlea. The steps include:

1. Sound Window
2. Frequency Mapping
3. Waveform Analysis
4. Stimulation Waveform
5. Stimulation Rate
6. Current Sources

Sound window. The *sound window* is the range of acoustic input levels that are allowed into the implant's signal-processing system. In normal listeners, the dynamic range for hearing is as much as 120 dB (re: acoustic hearing threshold). In contrast, the electrical dynamic range for cochlear implant users is very narrow, ranging between 10 and 20 dB (re: electrical hearing threshold). Consequently, a cochlear implant must transform the acoustic signal so that important information is retained and encoded within the narrow dynamic range of electrical hearing.

There are two methods by which HiRes handles the narrow electrical dynamic range of implant users—(1) by widening the dynamic range of sound entering the system (Input Dynamic Range, or IDR) and (2) by using a dual-action automatic gain control (AGCII). The microphone in the system has an instantaneous input dynamic range of 84 dB, which can easily accommodate the peaks and valleys of speech that can vary as much as 50 dB. Within the system, the IDR is programmable from 20-80 dB. If the IDR is too narrow, sounds below the IDR are discarded while sounds above the IDR are clipped or compressed for a given processor setting. The AGCII consists of an AGC amplifier whose gain is determined by two voltage controls. The slow-acting control has a compression threshold of 57 dB SPL with an attack time of 325 ms and a release time of 1000 ms. The second control is fast-acting and has a higher compression threshold of 65 dB SPL with an attack time of <0.6 ms and a release time of 8 ms. Typically, the slow-acting control determines the amplifier gain for everyday listening situations, while the fast-acting control activates in response to transient loud sounds.

Thus, the system is able to encode a wide range of speech intensities within a user's electrical dynamic range without requiring processor adjustments. The wider IDR enables HiRes users to hear speech spoken at many levels by different speakers from many distances. Figure 3 shows the effect of changing the IDR on word and sentence recognition (Allen and Donaldson 2000). For 70-dB input levels, IDR has no effect. For 60-dB input levels, consonant recognition begins to be affected. At 50 dB SPL, there is a greater effect of IDR on speech recognition for both consonants and sentences. In short, as the overall level of speech is reduced, a wider IDR is needed for improved speech recognition. Thus, a wide IDR is important for hearing speech in the real world where the levels can vary greatly depending, for example, upon the speaker's vocal effort or distance from the implanted listener.

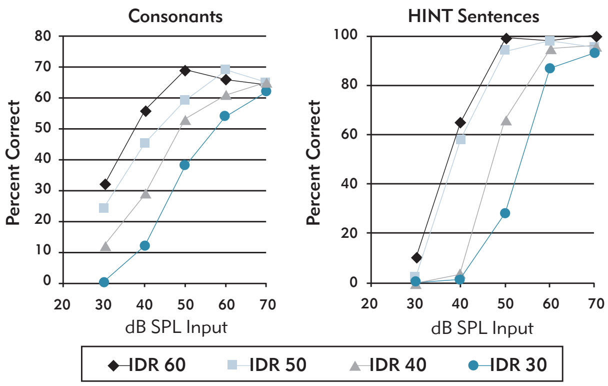

Figure 3. Consonant and sentence performance-intensity functions for four IDRs.

Frequency mapping. Once the acoustic signal has entered the implant system, it is digitized and sent through a set of bandpass filters that typically span the range of 250 to 8,000 Hz (normal listeners can hear between 20 and 16,000 Hz). In the HiResolution Bionic Ear system, there are 16 logarithmically spaced bandpass filters. The outputs from each of the 16 filters (after further processing, see below) are sent to the 16 respective stimulating contacts in the cochlea. Information from the lower frequency filters is sent to the apical contacts, while information from the higher frequency filters is sent to the basal contacts. In this way, spectral information is sent to different locations in the cochlea, thus mimicking the tonotopic mapping in a normal ear.

Waveform analysis. Frequency mapping provides essential information about the spectrum of the sound (that is, amplitude across frequency), however, temporal information in each spectral band also is important for speech understanding and music appreciation (see Shannon et al. 1995, Van Tassel et al. 1992). According to Rosen (1989), the temporal structure of speech can be classified into three categories based on the periodicity of the dominant temporal structure in each category. In his view, "envelope" cues are present from 2-50 Hz and provide prosody (stress, syllabification) information. "Periodicity" cues from 50-500 Hz indicate segmental information like consonant manner and voicing, and intonation. Temporal "fine structure" cues from 600-10,000 Hz carry information about consonant place and vowel quality.

Previous Advanced Bionics® systems were able to encode the temporal variations in the outputs of 8 filters out to only 200-400 Hz. However, HiRes employs a different algorithm to follow the temporal fluctuations up to 2800 Hz across 16 filters. Thus the system is able to replicate the timing information in the original signal several orders of magnitude faster than before, thereby providing the "fine structure" cues that can aid speech recognition, especially in noise, and music appreciation.

Once the temporal variations in the signal are extracted from each of the 16 filter outputs, the resulting waveforms are then used to modulate a pulse train. The modulated pulse trains are sent to the corresponding electrode contacts along the cochlea.

Stimulation waveform and rate, current sources. Although the HiResolution Bionic Ear system is able to deliver both pulsatile and analog waveforms to the electrode contacts, HiRes Sound currently uses biphasic pulses and monopolar coupling. The pulse widths can range from 11 to 229 microseconds per phase. The number of pulses per second (pps) per contact is a function of the pulse width, the number of channels used (1-16), and whether the pulses are delivered sequentially or to two contacts simultaneously (paired). The system has 16 independent current sources that allow for simultaneous stimulation of two or more electrode contacts. HiRes Sound adjusts the parameters to maximize the stimulation rate based upon the number of electrodes used and the pulse widths required by each patient. For 16 electrodes using sequential stimulation and narrow pulse widths, the rate is about 2900 pps per contact. For 16 electrodes using paired pulses and narrow pulse widths, the stimulation rate exceeds 5100 pps per contact.

It is the combination of fine temporal waveform representation and a fast stimulation rate that provides the high-resolution signal to the implant patient. Figure 4 shows the resulting electrode stimulation pattern for low-resolution temporal representation, and a slow stimulation rate. In contrast, Figure 5 shows the resulting electrode stimulation pattern for high-resolution temporal representation and a high stimulation rate. The stimulation pattern in Figure 5 results in the closest representation of the original signal waveform, which is preserved at the electrode contact.

HiRes sound processing, professionals, and implant candidates

For professionals, HiRes sound processing is best understood by appreciating its technical configuration and by having experience with patients using HiRes. The 16 electrode contacts, 16 independent current sources, dual-action AGC, and wide IDR are relevant. In addition, representation of the entire frequency spectrum and closer extraction of the temporal envelope, along with a fast stimulation rate, are important HiRes features. It could be said that HiRes combines the better parts of SAS and CIS, that is, full-spectrum signal analysis, more precise temporal waveform representation, and high-rate pulsatile stimulation.

When sharing this information with implant candidates or patients, simpler explanations of these signal-processing methods are needed. For some patients, the terms "newer," "faster," and "better" may suffice, along with confirmation that HiRes can be used with a behind-the-ear processor. For other patients, some explanation of envelope following and stimulation rate may be appropriate concepts to discuss. Both professionals and patients benefit from a description of the potential everyday benefits to be

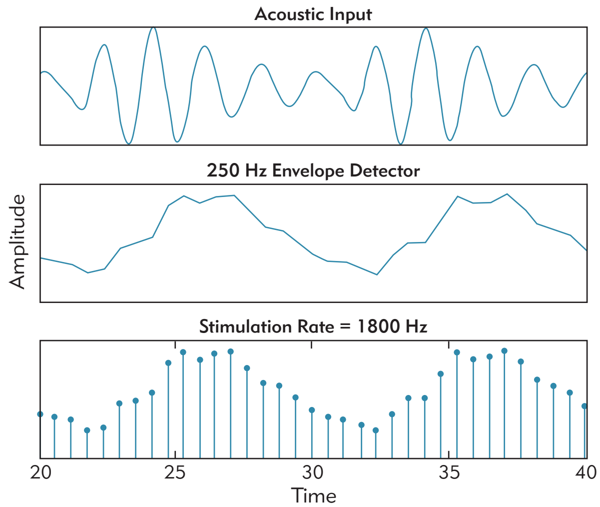

Figure 4. Acoustic stimulus waveform, temporal waveform extraction, and stimulation pulse pattern for 250-Hz temporal resolution and an 1800 pulses-per-second stimulation rate (low resolution-low rate).

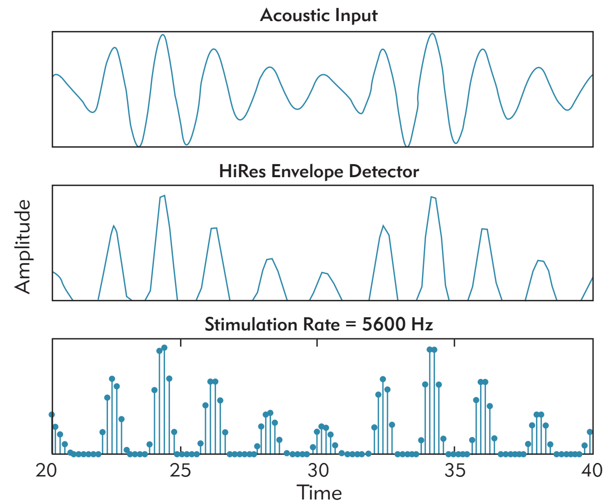

Figure 5. Acoustic stimulus waveform, temporal waveform extraction, and stimulation pulse pattern for HiRes temporal resolution and a 5600 pulses-per-second stimulation rate (high resolution-high rate).

gained from HiRes sound processing. Clinical study results indicate that patients who have used both conventional sound processing (e.g., SAS or CIS) and HiRes Sound show improved speech perception with HiRes, and that most patients prefer HiRes to conventional strategies (Koch et al. 2004).

In summary, the advanced signal-processing capabilities of the HiResolution Bionic Ear provide a high-fidelity representation of the acoustic signal to the cochlear implant user. Speech perception benefit may improve compared to conventional sound processing, and generally patients prefer HiRes Sound to older strategies. We cannot ignore the fact that outcomes with cochlear implants are also influenced by patient variables (e.g., for adults, length of auditory deprivation, residual hearing; for children, age at implantation, communication mode). However, we also cannot ignore the contribution that improved technology has on patient performance. Improving the electrical representation of the original acoustic signal can only improve the chance for better encoding by the patient, irrespective of the status of the patient's hearing and communication history. The implementation of HiRes sound processing takes an important step towards improved signal representation, and therefore, ultimately patient performance. Studies are underway to explore further the additional sound processing capabilities offered by the HiResolution Bionic Ear system.

Acknowledgements

This article is based upon a presentation at the 9th Symposium on Cochlear Implants in Children, Washington, DC, April 24-26, 2003. The assistance provided by Dawn Koch and Mike Faltys of Advanced Bionics is appreciated.

REFERENCES

Allen SL, Donaldson GS. (2000) unpublished data.

Koch DB, Osberger MJ, Segel P, Kessler D. (2004) HiResolution and conventional sound processing in the HiResolution Bionic Ear: using appropriate outcome measures to assess speech-recognition ability. Audiol Neuro-Otol (in press).

Rosen S. (1989) Temporal information in speech and its relevance for cochlear implants. In Cochlear Implant: Acquisitions and Controversies, ed. B. Fraysse, N. Couchard, pp. 3-26.

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

Van Tasell DF, Greenfield DG, Logemann JJ, Nelson DA. (1992) Temporal cues for consonant recognition: training, talker generalization, and use in evaluation of cochlear implants. J Acoust Soc Am 92(3):1248-1257.