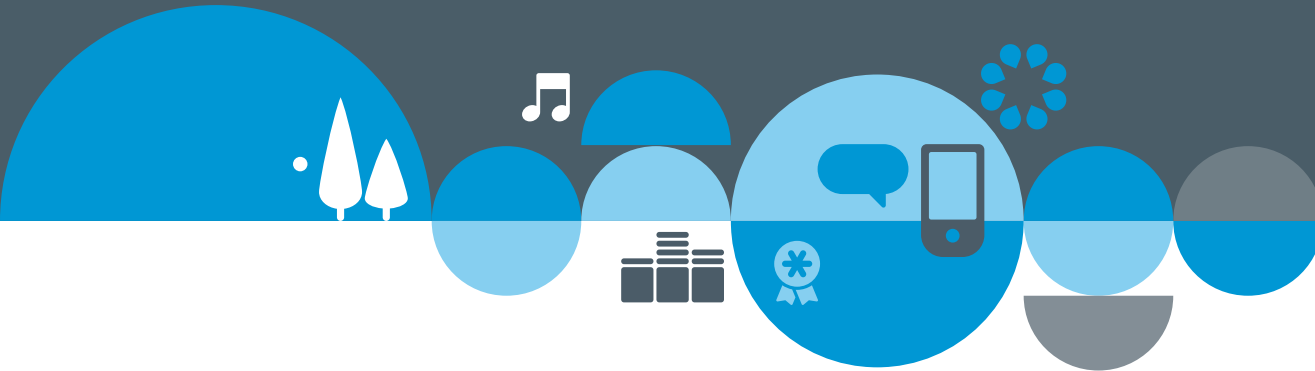


Sonic Spotlight



Speech Variable Processing: Delivering Natural Sound and Speech Clarity

While present-day digital signal processing (DSP) strategies almost exclusively rely on some method of compression or automatic gain control, implementations within hearing instruments still vary greatly, depending on manufacturer or even within families from top-end to entry-level performance categories.

Historically, early digital signal processing in hearing instruments provided a small number of channels to shape gain curves and allow different compression ratios at different frequencies. Today some systems partition the incoming signal into upwards of twenty channels before averaging the signal together over a segmented time span. Whereas some DSP platforms can be too simplistic, overly complex processing schemes can introduce artifacts, e.g. unwanted time delays, or spectral smearing – not just for consonants, but for vowel identification too (Souza, et al. 2012).

The Speech Variable Processing (SVP) strategy from Sonic effectively and efficiently addresses the need for optimum time and frequency resolution in order to provide exceptional speech clarity and natural sound. This document will describe the benefit of this approach and explain the basic concepts behind its design.

The Cochlear Amplifier

Before we describe how the Speech Variable Processing system functions, it is important to understand the role the cochlea plays in how we hear. When the cochlea is healthy, the outer hair cells amplify soft sounds and control very loud sounds – a process known as the “cochlear amplifier”. The theory was first promoted by Gold in 1948 and has been the subject of intense research over the following decades. As generally understood today, the process results from “[...] an active mechanical response in the cochlea that amplifies low-level and compresses high-level basilar membrane displacements” (Canlon, 2010).

Damage to the outer hair cells affects the cochlear amplifier’s ability to properly amplify sounds, resulting in sensorineural hearing loss. Clinical symptoms of sensorineural hearing loss are a significant loss of hearing for low-intensity sounds, some loss of sensitivity for moderate-intensity sounds, and near normal perception for high-intensity sounds. In other words, sensorineural hearing loss manifests in a non-linear way – the amount of loss is not the same at all sound levels.

The introduction of wide dynamic range compression (WDRC), which amplifies low-intensity sounds more than high-intensity sounds, provided a major leap forward in addressing sensorineural hearing loss.

At a very basic level, WDRC tries to replicate the cochlear amplifier model; however, WDRC alone does not take into account other aspects of how the cochlea processes sound.

For example, immediately upon sound reception, the cochlea functions as a signal analyzer, operating in narrow frequency-specific regions. This instantaneous physiological process along the tonotopically-arranged basilar membrane results in sharp frequency selectivity (Fettiplace, 2010). In order to replicate this functionality in a way that sounds natural, a processing system must be able to very quickly analyze incoming sounds, process the frequency-specific content, and deliver output at the appropriate level for the loss with as little delay as possible. Although not fully understood, the cochlear amplifier’s synergistic process has inspired the design of digital amplification.

The Mechanics of Speech Variable Processing

Speech Variable Processing is a very fast-acting WDRC system designed to replicate the cochlear amplifier model to keep sounds natural and speech clear. While many processing systems can be described in a similar manner, the implementation specifics are what make Speech Variable Processing a superior system.

Speed is key to accuracy

The cochlea’s ability to quickly analyze and internally adjust amplification of sound is impressive, and the key to our natural perception of sound volume and quality. Brownell (2010) explains that action potentials of the cochlear amplifier are responsible for encoding the temporal precision from environmental sounds. The SVP digital processing system also aims to preserve this valuable temporal information. The concept of speed can be looked at several different ways.

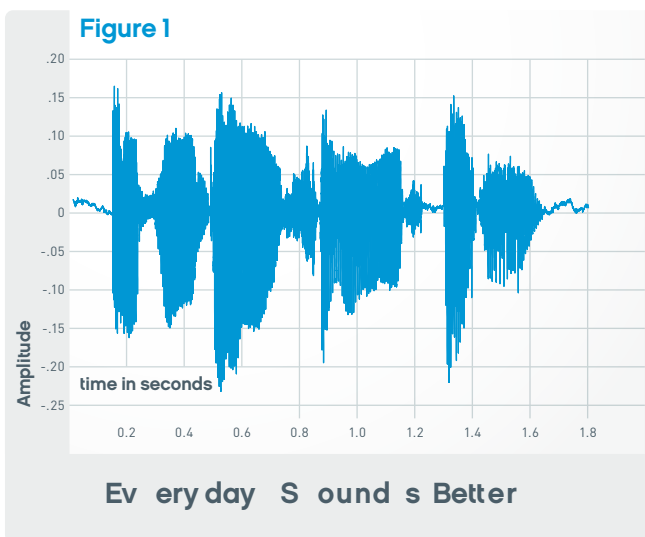


Figure 1: Sound pressure level differences over time for the incoming signal, “Everyday Sounds Better”.

First, it is imperative to quickly and accurately measure the input level of the incoming signal. Input levels are rarely static – they change rapidly (Fig. 1). This is especially true of speech, with individual phonemes occurring on average ten times per second. Phonemes can exhibit dramatic intensity differences, which if not measured quickly and correctly will not be amplified correctly either (Fig. 2). Second, once the level of the input is known, the required amplification must be applied. Again, speed is of the essence – the system must be able to apply the correct amount of amplification for just the right amount of time at just the right frequency. If a system cannot apply the correct amount of gain quickly enough, the amplification will lack the resolution necessary for sound to be natural and speech to be clear.

Speech Variable Processing analyzes the incoming signal and adjusts gain thousands of times per second. This remarkable amount of computational power allows signal processing that swiftly follows changes in sound pressure level, thus enhancing speech intelligibility and creating a natural sound. This particular DSP approach is able to process audio events without the constraints of slower time frames that are inherent in other signal processing schemes. Without those slower segmented time frames in the way, SVP is able to adapt gain smoothly and continuously, without disruption to the original signal. Only in this way is it possible to react without delay – exactly at the point in time when a sudden audio event occurs. And only in this way is it possible to follow the changes in sound pressure level during a speech utterance, which may include differences of up to 30 dB between vowels and consonants. This also helps to boost consonants, the part of speech that usually creates confusion among hearing impaired people, because each incoming signal rapidly receives the required amount of amplification (Fig. 3).

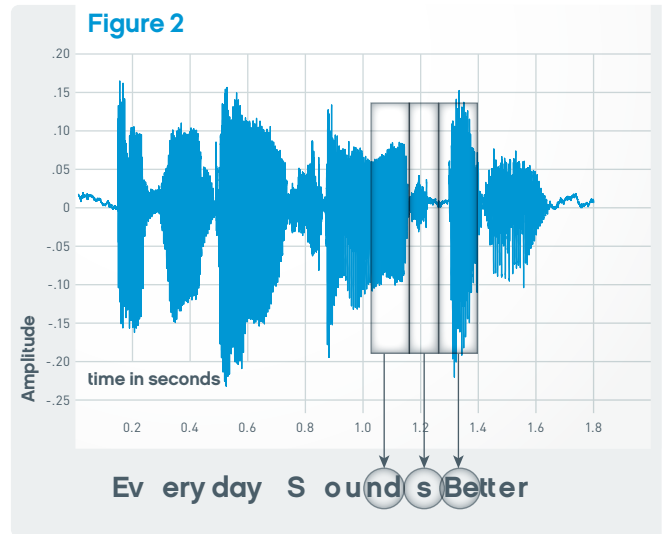


Figure 2: The intensity levels between the phonemes /nd/, /s/, and /B/ are very different. Speech Variable Processing's unique temporal resolution characteristics are able to capture and amplify each individual phoneme precisely, including the softest parts of speech.



Figure 3: Systems with fast attack and slow release times (orange) typically apply too much or too little gain because they do not react quickly enough to changes in sound pressure levels. With its very fast attack and release times, Speech Variable Processing (blue) is able to accurately measure and compensate for sudden changes in sound pressure levels.

The importance of frequency contrast in speech clarity

If speed is key to accuracy, then frequency contrast is key to clarity. Frequency contrast is the difference in intensity across frequencies of an incoming sound. Correctly identifying frequency contrasts play a large role in speech clarity – the frequency contrast between individual phonemes provides clues for listeners to tell different phonemes apart.

In a healthy cochlea, sound is amplified to maintain natural level differences across frequencies. With a sensorineural hearing loss, the ability to amplify correctly at different frequencies is diminished, resulting in a loss of frequency contrast. The processing system should be able to preserve the contrast inherent in the incoming signal to maintain the cues that increase speech identification and speech clarity.

Most traditional multichannel compression systems (1) split the acoustic signal into separate frequency bands, (2) measure the signal at separate levels for each

band, and (3) compress each band according to its individual measurement and prescribed gain setting. In doing so, these systems flatten the overall spectrum of sounds they process, losing the frequency contrast in the process.

Speech Variable Processing overcomes this loss of frequency contrast by measuring and applying gain to the wideband acoustic signal. Because the system is never measured and adjusted in separate frequency regions, the deleterious summing effect of applying gain in multiple bands is avoided, thereby preserving the contrast inherent in the original signal (Fig. 4). This also becomes especially important when preserving the peak-to-valley ratio that is required for vowel identification with hearing-impaired persons. Traditional multichannel compression systems have been shown to flatten the spectral detail of vowel formants as channel numbers increase, which can significantly decrease vowel identification performance for certain types of losses [Souza, 2012; Bor et al, 2008].

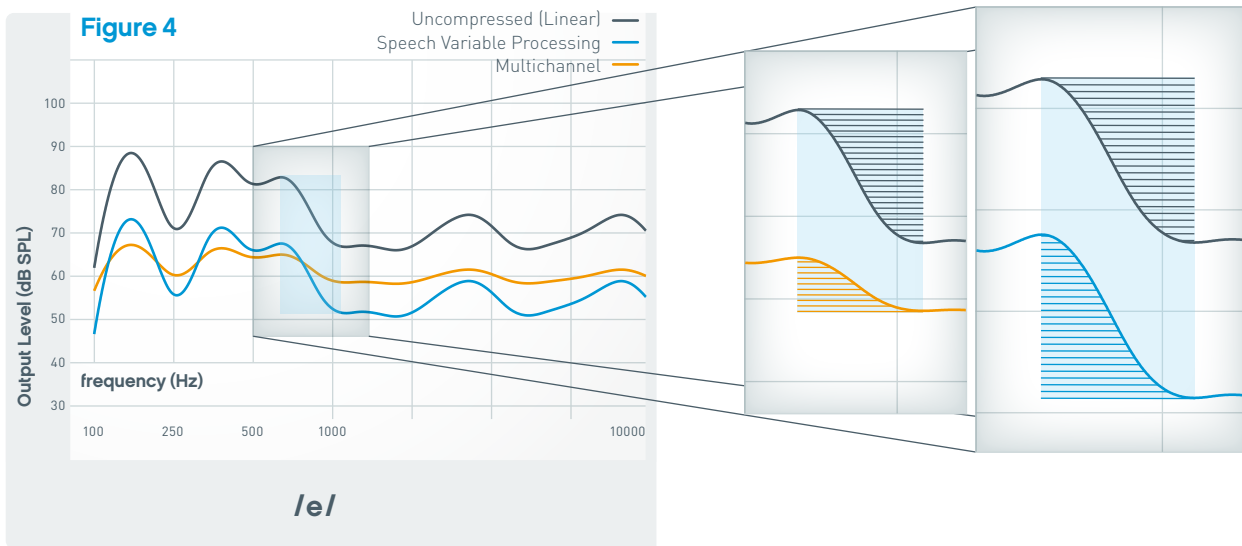


Figure 4: Illustration of the frequency contrast for the vowel sound /e/. Speech Variable Processing preserves the peak-to-valley ratio of the input signal, important for maintaining the frequency contrast of speech.

Speech Variable Processing – building on the Sonic legacy

The strength of SVP is that it fulfills the requirement of both time and frequency resolution needed for superior sound processing. It provides sufficient time resolution, reacting immediately to changes in sound pressure level, both when levels suddenly increase as well as suddenly drop. It also maintains the frequency contrast of speech, by upholding the peak-to-valley ratios present in the incoming signal.

Sonic has always believed the best signal processing could be achieved by using the cochlear amplifier as

inspiration. Speech Variable Processing is the next logical step in the evolution of Sonic's processing systems designed to make sounds audible, natural, and clear.

- Temporal accuracy is achieved through very fast processing speeds
- Speech clarity is enhanced by employing processing strategies that preserve frequency contrast

For a demonstration or to learn more, please contact your local Sonic provider.

References

Bor, S., Souza, P. & Wright, R. [2008]. Multichannel compression: Effects of reduced spectral contrast on vowel identification. *Journal of Speech, Language, and Hearing Research*, 51, 1315-1327.

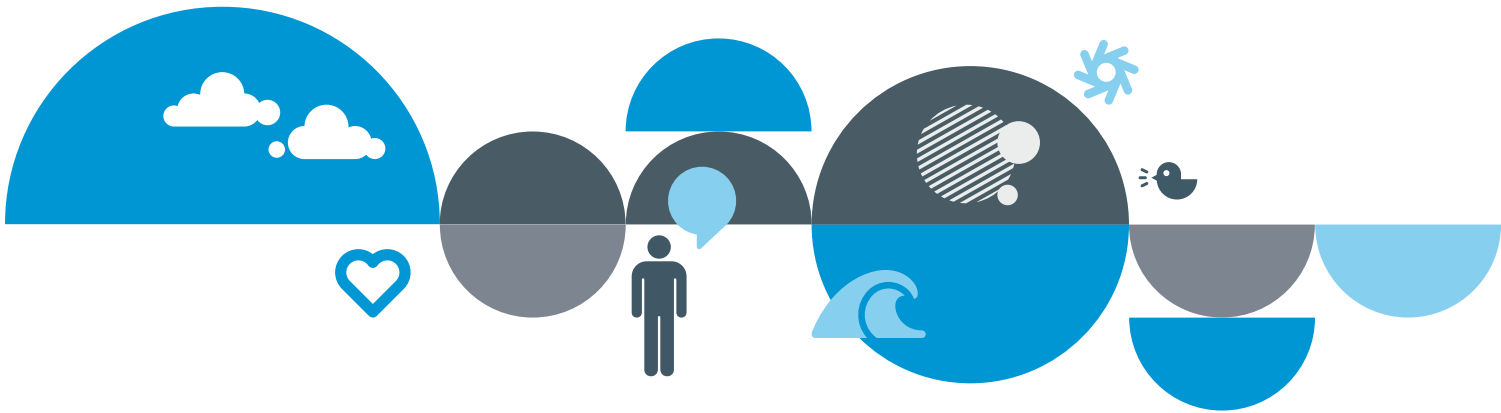
Brownell, W. [2010]. The remarkable cochlear amplifier. *Hearing Research* 266, 1-17; Section: "Membrane-based amplification in hearing."

Canlon, B. [2010]. The remarkable cochlear amplifier. *Hearing Research* 266, 1-17; Section: "Introduction."

Fettiplace, R., Hackney, C. [2010]. The remarkable cochlear amplifier. *Hearing Research* 266, 1-17; Section: "The origin of the cochlear amplifier."

Gold, T. (1948). *Hearing II. The physical basis of the action of the cochlea.* Proc. Royal Soc Lond. Biol Sci 135, 492-498.

Souza, P., Wright, R., Bor, S. [2012]. Consequences of Broad Auditory Filters for Identification of Multichannel-Compressed Vowels. *Journal of Speech, Language, Hearing Research*; 55; 474-486.



Sonic Innovations, Inc.
2501 Cottontail Lane
Somerset, NJ 08873 USA
+1 888 423 7834

Sonic AG
Morgenstrasse 131B
3018 Bern, Switzerland
+41 31 560 21 21