

Compression in Hearing Aids: Why Fast Multichannel Processing Systems Work Well

Villchur on the controversies surrounding compression in regard to recruitment and speech intelligibility in noise

References in the academic literature maintain that flattening of the speech envelope by fast multichannel amplitude compression reduces speech intelligibility—references that typically ignore the counteracting effect of recruitment. Meanwhile, almost every modern hearing aid uses some version of this compression. This contradiction needs to be resolved.

By Edgar Villchur, MEd

Fast multichannel amplitude compression in hearing aids reduces short-term amplitude contrasts among elements of speech and flattens the speech envelope. Critics of this processing claim that short-term amplitude contrasts are important cues to speech recognition, and that the processing reduces speech intelligibility. This article points out:

- 1) Listeners can only be aware of contrasts in physical amplitude as loudness contrasts; the reduction of short-term amplitude contrasts by fast compression is designed to compensate for the exaggerated loudness contrasts created by recruitment and to restore these loudness contrasts to normal.
- 2) High-frequency emphasis, like compression, reduces short-term amplitude contrasts and flattens the speech envelope, but there is general agreement (among critics as well as advocates of fast compression) that such emphasis increases intelligibility for people with accentuated high-frequency loss.

Reduction of Amplitude Contrasts and Criticisms of Fast WDRC

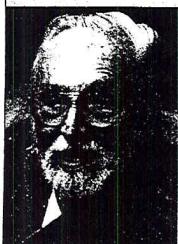
Recruitment creates a progressive recovery from an elevated hearing threshold toward normal loudness response as

the sound level increases, until at high levels the hearing-impaired listener may hear sounds at normal or near-normal loudness. Fast wide-dynamic-range compression (WDRC) in hearing aids increases the relative gain applied to weak speech elements so they can be amplified to intelligible levels without having to overamplify (from the point of view of the listener with recruitment) high-amplitude speech elements to uncomfortable levels.

But critics of this processing say it has an effect on speech, inherent in the processing, that degrades intelligibility. That effect is the reduction of short-term amplitude contrasts among elements of speech, contrasts that are claimed to be important cues to speech recognition. Some critics give compression credit for reducing the dynamic range of speech to fit the reduced dynamic range of hearing resulting from recruitment, but they say the advantage is mitigated or erased by the speech-degrading effect of the processing.

Plomp¹ referred to this reduction of amplitude contrasts as a reduction of the modulation transfer function (the envelope of speech looks like an amplitude-modulated carrier); Drullman et al² called it a reduction of temporal fluctuations; Kuk³ called it temporal distortion; Goldstein⁴ called it a loss of temporal modulation; Moore⁵ called it distortion of the temporal envelope; Olsen⁶ called it a lack of preservation of the

temporal envelope contrast; and Ronan et al⁷ called it short-term alterations of the overall spectral envelope. The words “temporal,” “envelope,” or “modulation” limit the application of the above terms to contrasts of successive rather than simultaneous speech elements



Edgar Villchur, MEd, is president of the nonprofit laboratory Hearing Aid Research, Woodstock, NY. His 1973 paper in the *Journal of the Acoustical Society of America* is considered one of the seminal papers on compression in hearing aids, and he helped develop one of the first hearing aid systems that combined compression and equalization, manufactured by ReSound. Villchur is also the inventor of the acoustic suspension loudspeaker (generally considered in the stereo industry as one of the five most important speaker designs since the Rice-Kellog dynamic speaker of 1925) and the dome tweeter. He has served as a visiting scientist at MIT, and is the author of numerous articles and books on acoustics.

(the latter as in vowel formants), but this distinction may not have been intended.

Critics of fast multichannel compression have also claimed that compression creates an undesirable interaction between fluctuating interference and the target signal when the two are in the same frequency channel.

The purpose of this article is to respond to the theoretical objections referred to above; the usefulness of fast multichannel compression in hearing aids must be proved or disproved by experimental data.

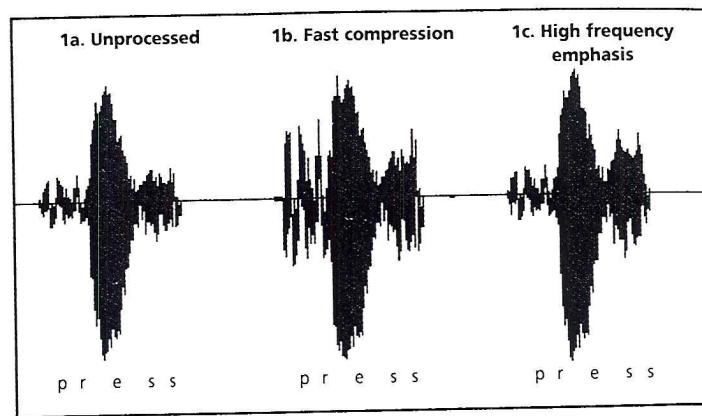
Intensity Versus Loudness Contrasts: Compression Restores Normal Loudness Relative to Recruitment

Viewing recruitment as a loudness expander. Fast compression *does* reduce short-term amplitude contrasts among elements of speech, whether the elements are successive or, when they are in different frequency channels of the compressor, simultaneous. That is what compression is designed to do.

But listeners can only be aware of contrasts in physical amplitude as *loudness contrasts*, and a listener with recruitment hears amplitude contrasts with exaggerated contrasts in loudness, as though listening through an *electronic expander*. In fact, Steinberg and Gardner,⁸ who first analyzed recruitment in 1937, described its effect as analogous to that of an expander.

Today, we know that outer hair cells compress the signal to fit the limited dynamic-range capacity of the inner hair cells. When this physiological compression is impaired, the signal reaching the inner hair cells is less compressed than in the normal ear. The signal is thus expanded relative to the normal compressed signal, creating recruitment.

Steinberg and Gardner⁸ used subjects with unilateral hearing impairment and recruitment to demonstrate the similarity between recruitment and expansion. They measured the required increase in the SPL of a signal presented to a subject's recruiting ear over the SPL of a signal presented to his/her normal ear for the two signals to be equal in loudness, and repeated the measurement over a range of levels. The difference between signal SPLs that produced the same loudness in each ear became less as the level increased; for some subjects the difference disappeared at high



FIGURES 1a-c. Left (1a): Unprocessed wave envelope of the spoken word "press." Middle (1b): Wave envelope after the signal has been processed by fast compression. Right (1c): Wave envelope after the signal has been processed by high-frequency emphasis. The amplitude contrast between the low frequency vowel /e/ and high frequency consonant /s/ has been reduced as in the compressed signal, but the contrast between the vowel and the low frequency consonants /p/ and /r/ is unaffected.

levels, so that a given high-level signal produced the same loudness in the subject's impaired ear as in his/her normal ear.

A demonstration of the similarity between recruitment and expansion was also made by Villchur,⁹ who processed speech with a bank of expanders to simulate the individual recruitment characteristics of unilaterally impaired subjects. The subjects judged the simulation presented to their normal-hearing ears as similar or very similar to the real thing in their impaired ears. However, when expansion was removed from the processed signal the sounds were judged as very different. Duchnowski and Zurek¹⁰ reported that the speech-test errors made by normal-hearing subjects listening to an expander simulation of recruitment were similar to the errors made by subjects with actual recruitment.

Figure 1a is a graphic record of the envelope of the spoken word "press," with the vertical axis representing amplitude in sound-pressure units and the horizontal axis representing time. Fast compression (Figure 1b) applied to this signal amplifies the low-amplitude sections of the signal more than the high-amplitude section and reduces the amplitude contrasts between them.

Addressing the expansion characteristics of recruitment with compression. The speech envelope can also be plotted on a graph in which the scale of the vertical axis is in *loudness units* rather than amplitude units; the graph then represents the way the amplitude contrasts are *perceived*. If the sound heard by a person with recruitment is plotted on such a graph, the weaker elements will be depressed and the loudness envelope will show greater contrasts than the contrasts heard by a normal-hearing listener. When fast compression is applied to the

physical signal, and the reduction of amplitude contrasts by compression is reasonably matched to the exaggeration of loudness contrasts by recruitment, the loudness contrasts will be restored to (or toward) the contrasts heard by a person with normal hearing. With ideal matching of compression and recruitment, the loudness envelope will be the same as the one perceived by a normal listener.

Amplitude/loudness contrasts as elements in speech recognition. Unless it can be shown that listeners with recruitment rely on enhanced loudness contrasts for speech recognition — to make up for reduced frequency resolution, for example, as Plomp¹ and others have suggested — a reduction of amplitude contrasts by compression that matches the recruitment should not be expected to decrease speech recognition. If hearing-impaired listeners *do* rely on enhanced loudness contrasts for speech cues, that would imply *recruitment can ameliorate their impairment*, as Fowler¹¹ suggested. (Specifically, Fowler said recruitment reduces the effect of hearing loss on speech recognition when the loss is less than 40 dB.) Evidence of that effect has not been reported.

Licklider and Pollack¹² reported that speech retained high intelligibility for normal listeners after all-pass amplitude contrasts had been eliminated at one stage of processing by infinite peak clipping. They concluded that intelligibility was retained in the pattern of time-axis crossings, and that "variations in intensity from moment to moment appear not to be basic cues for the recognition of words."¹² Plomp came to the opposite conclusion. He said, "intensity contrasts in this spectro-temporal pattern are the major vehicle of the speech information."¹¹ In any event, to the extent that amplitude contrasts provide cues to speech recognition, these cues are perceived as loudness contrasts, and when amplitude contrasts are reduced by compression properly adjusted to the recruitment of the listener, loudness contrasts are restored to normal.

Weaker elements of speech. The weak elements in some of the amplitude contrasts of speech may remain below the hear-

ing threshold of a listener with recruitment after linear amplification to his/her preferred overall speech level. This can occur with words that have strong vowels and weak consonants, like "both" or "laugh." Maintaining the integrity of the amplitude contrasts while maintaining the listener's preferred overall speech level would require keeping the weak elements inaudible.

These weak speech elements can be amplified to the desired loudness of a listener with recruitment in two ways:

- 1) By turning up the gain of a constant-gain amplifier from its setting at the listener's preferred speech level — that is, by increasing the gain for all parts of the speech for the sake of that part of the speech that needs extra gain, or
- 2) By using a variable-gain amplifier.

The constant-gain amplifier will maintain the integrity of amplitude contrasts and of the speech envelope; however, it will also maintain the exaggeration of loudness contrasts, and it will overamplify high-level speech elements. The variable-gain ampli-

er will violate the integrity of amplitude contrasts but correct the exaggeration of loudness contrasts; it will amplify low-level speech elements to the necessary loudness without overamplifying high-level sound, as though the listener turned up a manual volume control when weak sounds appeared. For a person with recruitment, the hearing aid can maintain integrity for the physical relation between the SPLs of speech elements, or integrity for their relative loudness, but not for both.

Number of channels. Plomp and others said that fast compression with a very large number of frequency channels and very large compression ratios for each channel will reduce speech to a "stationary sound without any structure."¹ Compression reduces the difference in dB between the amplitudes of two successive signals in the same compressor channel by a numerical divisor called the *compression ratio* (CR), assuming the compressor attack and release times are shorter than the time between signals. The difference between signal SPLs x and y after compression can thus be expressed as $(x \text{ dB} - y \text{ dB})/\text{CR}$. An infinite compression ratio and

an infinite number of channels *would* make all successive signals equal in amplitude.

Whether or not intelligibility can be retained after the elimination of amplitude contrasts, as Licklider and Pollack¹² reported, Plomp's prediction does *not* imply that fast multichannel compression with lower compression ratios and fewer channels puts a listener with recruitment on the path of progressively decreasing speech recognition. Swallowing the contents of a bottle of aspirin will make one sick, but that doesn't mean two aspirin will not relieve a headache.

High-frequency Emphasis, as Well as Fast Compression, Flattens the Speech Envelope

The amplitude contrasts in speech often occur between elements in different frequency regions, as in the contrast between the strong low-frequency vowel /e/ and weaker high-frequency consonant /s/ of Figure 1a. Accentuated high frequency hearing loss, like recruitment, will reduce the loudness of this consonant relative to that of the vowel, and like recruitment it will exaggerate the loudness con-

trast between the two.

It is almost universal practice to compensate for increased high frequency hearing loss by high frequency emphasis in the hearing aid. This emphasis, like compression, will increase the relative gain for the high frequency consonant /s/ in Figure 1a, and will reduce the amplitude contrast between the vowel and the consonant, as shown in Figure 1c. As with compression, high frequency emphasis violates the integrity of the amplitude contrast between these two elements of speech, but reduces the exaggerated loudness contrast between them. It amplifies the high frequency elements of speech—elements that are heard only faintly or not at all by a person with more severe high frequency loss—to intelligible levels without overamplifying low-frequency, high-amplitude elements. High frequency emphasis varies gain with frequency, while compression varies gain with level, but each reduces short-term amplitude contrasts and flattens the speech envelope.

If the reduction of short-term amplitude contrasts reduces speech cues and degrades intelligibility for hearing-impaired people, the same degradation of intelligibility ought to occur whether the amplitude contrasts have been reduced by the compression illustrated in Figure 1b or by the high-frequency emphasis illustrated in Figure 1c; the envelopes of the vowel and high frequency consonant in these two figures are essentially the same. But there is general recognition, by both critics and advocates of fast compression, that high-frequency emphasis increases speech intelligibility for persons with increased high frequency hearing loss.

Most people with cochlear hearing impairment have both recruitment and increased high frequency loss. The typical compression hearing aid is designed to compensate for both. Compression makes the frequency response level-dependent.

Attack and Release Times

The reason for choosing fast-acting rather than slow-acting compression is that the speech elements with contrasting amplitudes—on which the compressor must act successively—may be close together in time, as within the word “press” of Figure 1. A possible negative effect of short attack and release times in a compressor (less than 50 msec release time would be considered short) is that the quiet intervals between successive sounds are filled in at the edges by fast compression. This slows down the decay of the first sound (note the slight elongation of the decay of the vowel in Figure 1b), and the attack of the following

sound. (Author’s note: Fast-acting compression is sometimes called “syllabic compression,” implying that the attack and release times are short enough for the compression to act separately on successive syllables. The term does not, however, clearly imply separate compressor action on phonemes *within* a syllable. “Phonemic compression” would be a more comprehensive term.)

The reduced time intervals are no shorter than the intervals heard by a normal listener; recruitment lengthens silent

Most people with cochlear hearing impairment have both recruitment and increased high frequency loss. The typical compression hearing aid is designed to compensate for both. Compression makes the frequency response level-dependent.

intervals by speeding up the perceived attack and decay of sounds at the beginning and the end of the interval, and compression counters that effect. But it is possible that hearing-impaired people, particularly those with severe-to-profound loss, benefit from a demarcation between sounds that is sharper than that required by normal-hearing listeners, and thus from longer attack and release times.

Critics of fast compression often point out that slow compression adjusts overall speech levels to reduce long-term amplitude contrasts only, such as contrasts that occur among different voices or passages of speech or at different distances between talker and listener. The rationale for slow compressor action is that it preserves the short-term amplitude variations of the speech envelope. Indeed, slow compression does achieve this. However, preserving the short-term amplitude envelope allows details of the loudness envelope—which is what the listener hears—to remain distorted by recruitment.

The Just-Noticeable Difference for Intensity

Plomp¹ recognized that the reduction of amplitude contrasts by compression was opposite to the effect of recruitment. He wrote: “Strictly, the conclusions [on the effect of compression on speech recognition] are only justified for normal-hearing listeners.”¹

But he said his criticism of compression still applied because recruitment, contrary to expectations, fails to reduce the *just-noticeable difference* (JND, also called the *difference limen* or DL) for intensity, and compression therefore fails to create normal JNDs. The amplitude difference between two successive sounds that are one JND apart before compression is reduced by compression, so that detection of the difference after compression requires a larger pre-compression difference in SPL. Plomp implied that, since compression does not create normal JNDs, it will not restore normal loudness contrasts for larger amplitude differences. Further, he said a normal JND is important, in itself, because a listener cannot hear amplitude differences in speech that are smaller than his/her JND.

The amplitude JND of a hearing-impaired person with recruitment is typically normal, so that compression increases the JND to greater than normal. But that is not relevant to the effect of compression on the exaggerated loudness contrasts of larger amplitude intervals, which are reduced toward normal by compression.

Hellman et al¹³ measured amplitude JNDs in subjects whose normal loudness functions (ie, a curve that plots the relation between loudness and amplitude over a range of intensities) had been changed by masking. The slope of the loudness function was unrelated to the size of the JND, which is evidence that the size of JNDs does not affect the perception of larger amplitude differences.

I am not aware of any study showing that just-detectible amplitude contrasts in speech are significant to speech recognition.

Noise Reduction in Recording

Prior to digital recording, professional tape recordings of music or speech were made with a compression/expansion noise-reduction system. The signal was first processed by a fast multichannel compressor whose characteristics were very similar to those typically used in hearing aids. The compressor provided extra gain for weaker elements of the signal so that these elements could be recorded at levels well above the tape hiss.

When the signal was played back, it was processed by an expander whose characteristics were the mirror image of the compressor. The expander restored normal amplitude relations to the signal but attenuated the tape hiss, which had not been subject to the original compression. Noise-reduction processors were made by Dolby, dbx, and others.

Finding fault with fast multichannel

compression in hearing aids because it reduces amplitude contrasts and flattens the speech envelope is analogous to finding those same faults in the compression half of a Dolby system used without the expansion half. The hearing-aid/human-cochlea system through which a signal passes before it is perceived by a listener with recruitment cannot be evaluated without counting the effective expansion of the recruitment as an integral part of the system. Similarly, a hearing aid with high frequency emphasis cannot be evaluated without counting the listener's high frequency loss.

Compression and Noise

Plomp wrote that fast multichannel compression in the presence of fluctuating interference "will be still more detrimental because, for each frequency channel, the compression circuit will be equally sensitive to the noise fluctuations resulting in strong interactions."¹¹

The gain of a compressor channel is controlled by the level of the total input signal to the channel. When noise and the target signal are present in the same frequency channel, fluctuations in noise level affect the channel's momentary gain and therefore modulate the target signal. This effect, however, does not become significant until the noise level is equal to or higher than the level of the target signal—a circumstance in which the hearing aid is in any case not very useful. When the noise and target-signal levels are equal, total

RMS input to the compressor (the vector sum of signal and noise voltages, as shown in the sidebar) is increased 3 dB by the noise; when the noise is as little as 3 dB below the target signal, total input to the compressor is increased only 1.76 dB by the noise, and modulation of the signal by the noise is inaudible or next to inaudible.

However, high-intensity noise with a duration of only a few milliseconds will not be very loud because of its short duration and it will not have a significant masking effect; yet it will cause a compressor with a short attack time to reduce gain for both signal and noise. The clatter of plastic poker chips is such a stimulus. I have measured the sound of a single chip collision as exceeding 100 dB SPL for a period of less than 2 msec. The reduced compressor gain caused by that type of stimulus will recover at the compressor's slower release time, creating a "pumping" effect. My experience is that such pumping is not audible when the release time of the compressor is less than 50 msec. Another approach to the problem is to make the release time dependent on the duration of the transient sound.¹⁴

When the target signal is the predominant signal in the channel, strong target signals will drive the compressor to reduce gain for both signal and noise. In that case, the noise level does follow the gain variations of the compressor, but the effect is no different from the natural effect of unamplified high-amplitude target signals reducing the loudness of lower-amplitude noise by masking. The signal-to-noise ratio (SNR) in the compressor channel does not change when the target signal and noise are simultaneous, because at any one moment all signals in a channel are subject to the same gain.

When low-level noise occurs during quiet intervals of the target signal, compression will increase the noise level and reduce the SNR relative to target signals at other moments and in other channels. For listeners with recruitment, this SNR—like other amplitude contrasts—is reduced by the compression to, or toward, what it would be for normal-hearing listeners. (Recruitment acts like a built-in noise suppressor

when the noise is below the target signal, which increases the perceived SNR, and compression counteracts that effect.) Unlike the restoration of normal loudness contrasts in speech, the restoration by compression of a normal perceived SNR is a disadvantage; hearing-impaired listeners generally need a better SNR than do listeners with normal hearing. On the other hand, compression can increase the SNR between low-level speech elements in one channel and noise in other channels and at other moments; the compressor can increase gain for the low-level speech without affecting the noise level.

Fast compression has a dual effect on speech recognition in noise. Compression increases the gain for low-level elements of speech and restores redundant speech cues¹⁵ to hearing-impaired listeners. This provides a reserve of speech cues against the loss of cues to masking. Even though the noise has also been increased, the added cues increase the listener's ability to understand speech at a given noise level. Redundant cues are also restored to any undesired competing speech, but selective listening—the ability to pay attention to a particular message among simultaneous competing messages¹⁶—is a cognitive function, and can separate clear speech from clear interference more easily than muffled speech from muffled interference.

On the negative side, compression reduces the across-channel and across-time SNR when the interference level is below that of the desired signal, destroying speech cues.

The net effect of compression on speech recognition in noise depends on whether the positive effect of speech cues restored to audibility (ie, the increase in tolerance of the listener to noise) exceeds the effect of cues made inaudible by the increased noise. In a study on the effects of fast two-channel compression on hearing-impaired subjects in the presence of speech interference 10 dB below the target speech, Villchur¹⁷ reported that compression plus frequency-response shaping improved intelligibility scores significantly. This experiment was not, however, carried out with lower SNRs, and other investigators have reported opposite results.

Conclusion

The reduction of short-term amplitude contrasts among elements of speech by compression does not reduce speech cues for listeners with recruitment, unless such listeners require enhanced loudness contrasts as cues to speech recognition. That

What Happens When Noise Approaches or Equals the Level of Speech?

The gain of a compressor channel is affected by fluctuating noise in that same channel. However, this effect does not become significant until the noise equals or exceeds the target signal. When that is the case, the total signal voltage (e_t) = $\sqrt{e_s^2 + e_n^2}$

e_s = target-signal voltage;
 e_n = noise voltage;
 e_t = total voltage

When the noise level is only 3 dB below the target signal:

$e_t = \sqrt{e_s^2 + .708e_s^2}$
 $= \sqrt{e_s^2 + 0.5e_s^2}$
 $= \sqrt{1.5e_s^2}$
 $= 1.225e_s = 1.76 \text{ dB above } e_s$

At 2:1 compression, the resulting change in channel gain is only 0.9 dB—a change in gain that would be inaudible.

requirement has yet to be demonstrated. Properly programmed compression *restores*, rather than reduces, normal loudness contrasts for these listeners.

References in the academic literature to the supposed disadvantage of flattening the speech envelope—references that typically ignore the counteracting effect of recruitment—continue, while almost every modern hearing aid uses some version of the compression described above. This contradiction needs to be resolved. ▀

Acknowledgement

I would like to express my appreciation to Mead C. Killion, PhD, for his valuable contribution to this article, in particular for creating Figure 1.

References

1. Plomp R. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation transfer function. *J Acoust Soc Am*. 1988;83:2322-2327.
2. Drullman R, Festen JM, Plomp R. Effect of tem-

- poral envelope smearing on speech reception. *J Acoust Soc Am*. 1994;95(2):1053-1064.
3. Kuk F. Theoretical and practical considerations in compression hearing aids. *Trends Amplif*. 1996;s1(1).
4. Goldstein JL. Cochlear signal processing for compression and gain control extends dynamic range and preserves temporal modulation. Paper presented at: Second Biennial Hearing Aid Research and Development Conference; National Institutes of Health (NIH), Bethesda, Md; 1997.
5. Moore BCJ. Cochlear hearing impairment and the design of hearing aids. *Echoes*. 1998;8 (Fall 1998).
6. Olsen LO. Supra-threshold hearing loss and wide dynamic range compression. Vallingby, Sweden: Elanders Gotab; 1994:54.
7. Ronan D, Dix AK, Shah P, Braida LD. Integration across frequency bands for consonant identification. *J Acoust Soc Am*. 2004; 116(3):1749-1762.
8. Steinberg JC, Gardner MB. The dependence of hearing impairment on sound intensity. *J Acoust Soc Am*. 1937;9:11-23.
9. Villchur E. Simulation of the effect of recruitment on loudness relationships in speech. *J Acoust Soc Am*. 1974;56:1601-1611 [recording bound in with article].
10. Duchnowski P, Zurek PM. Villchur revisited: Another look at automatic gain control simulation of recruiting hearing loss. *J Acoust Soc Am*. 1995;98: 3170-3181.
11. Fowler EP. A simple method of measuring percentage of capacity for hearing speech. *Arch Otolaryngol*. 1942;36(6):874-890.
12. Licklider JCR, Pollack I. Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech. *J Acoust Soc Am*. 1948;20:42-51.
13. Hellman R, Scharf B, Teghtsoonian M, Teghtsoonian R. On the relation between the growth of loudness and the discrimination of intensity for pure tones. *J Acoust Soc Am*. 1987;82:448-453.
14. Killion MC. The K-AMP hearing aid: an attempt to present high fidelity for the hearing impaired. Recent developments in hearing instrument technology. In: Beilin J, Jensen GR, eds. *15th Danavox Symposium, Denmark*; 1993:167-229.
15. Coker CH. Speech as an error-resistant digital code. *J Acoust Soc Am*. 1974;55:476(A).
16. Broadbent DE. *Perception and Communication*. New York: Pergamon; 1958.
17. Villchur E. Signal processing to improve speech intelligibility in perceptive deafness. *J Acoust Soc Am*. 1973;53:1646-1657.

Correspondence can be addressed to HREditor@ascendmedia.com or Edgar Villchur at EVillchur@aol.com.