A New Technique for Quantifying Temporal Envelope Contrasts

Todd W. Fortune, Brian D. Woodruff, and David A. Preves

A new technique has been developed for precisely quantifying the temporal contrasts that exist between two sound samples. This technique is based on envelope subtraction, and generates an Envelope Difference Index that may be used to help clarify whether alteration of the natural speech envelope via amplification improves or degrades speech intelligibility. The Envelope Difference Index method may also be used to assess hearing aid saturation, and may have other applications as well. The technique is applicable whenever a precise quantification of the difference between two temporal envelopes is required, regardless of stimulus duration.

(Ear & Hearing 1994;15:93–99)

The influence of the temporal effects of amplification on the intelligibility of speech has been a topic of discussion for many years. Some researchers (Plomp, 1988; Moore, 1990) have maintained that multi-band amplitude compression circuits, particularly those with short time constants, can degrade speech intelligibility by reducing the natural temporal fluctuations of speech. Others have asserted that the temporal alteration of speech may be beneficial, if such alteration prevents loudness discomfort, increases the audibility of consonants (Villchur, 1989), or compensates for the effects of recruitment on temporal resolution (Moore, 1991).

The questions involved in this discussion are obviously quite complex, and much more research will be necessary before definitive conclusions about the relationship between amplification and speech intelligibility may be drawn. Conclusions will likely vary, however, depending on the method that is used to quantify the temporal changes to speech or other signals that occur as a result of amplification. Temporal effects have often been described in terms of the Modulation Transfer Function (MTF), the Speech Transmission Index (STI) or the simplified counterpart of the STI (RASTI), indices that relate speech intelligibility to the modulation depth of temporally fluctuating signals such as continuous discourse or amplitude-modulated noise (Houtgast & Steeneken, 1985). The MTF has been used to show how background noise and reverberation can reduce the modulation depth of speech. This temporal alteration is reflected in the STI, an MTF-based calculation of the overall signal-to-noise ratio that may be used to predict speech intelligibility. Although the MTF represents an effective method of analyzing temporal waveforms, it also has certain limitations. First, the MTF is based on average temporal effects measured over relatively long time intervals; it cannot be used to examine specific elements of speech. More could be learned about the temporal effects of amplification if it were possible to quantify temporal contrasts involving words, syllables, or even phonemes. The MTF technique is also limited in the types of signals that are suitable for analysis. It cannot be applied to nonspeech signals that often affect hearing aid performance, such as impulse noise. Finally, although the STI may be used to predict overall speech intelligibility for a particular speech material, it cannot be used to help explain individual behavioral responses. The following is a description of a new temporal analysis technique that has been designed to overcome these limitations.

The Envelope Difference Index

To quantify temporal changes caused, for example, by amplification, a technique for comparing the temporal envelopes of two acoustic signals has been devised. This technique is based on temporal envelope subtraction, and generates what will be referred to as the Envelope Difference Index (EDI). The essential elements of the technique are outlined in Figure 1. Panel A of the figure shows a digitized waveform of the syllable /it/. This waveform had been amplified by a Class D linear hearing aid, and was recorded in the ear of an experimental subject, using a probe microphone. Signal processing begins by taking the absolute value of this waveform (Panel B), which will allow the envelope to be described as a continuous, positive-amplitude function, relative to direct current (DC). The envelope of this waveform is derived by digital low-pass filtering (Panel C). The waveform shown was filtered at 20 Hz, although cutoff frequencies as high as 100 Hz may be used. Once the envelope of this aided waveform has been derived, it may be compared with the unaided envelope of the same syllable, obtained by the same

Argosy Electronics, Eden Prairie, Minnesota

0196/0202/94/1501-0093$3.00/0 • Ear & Hearing •

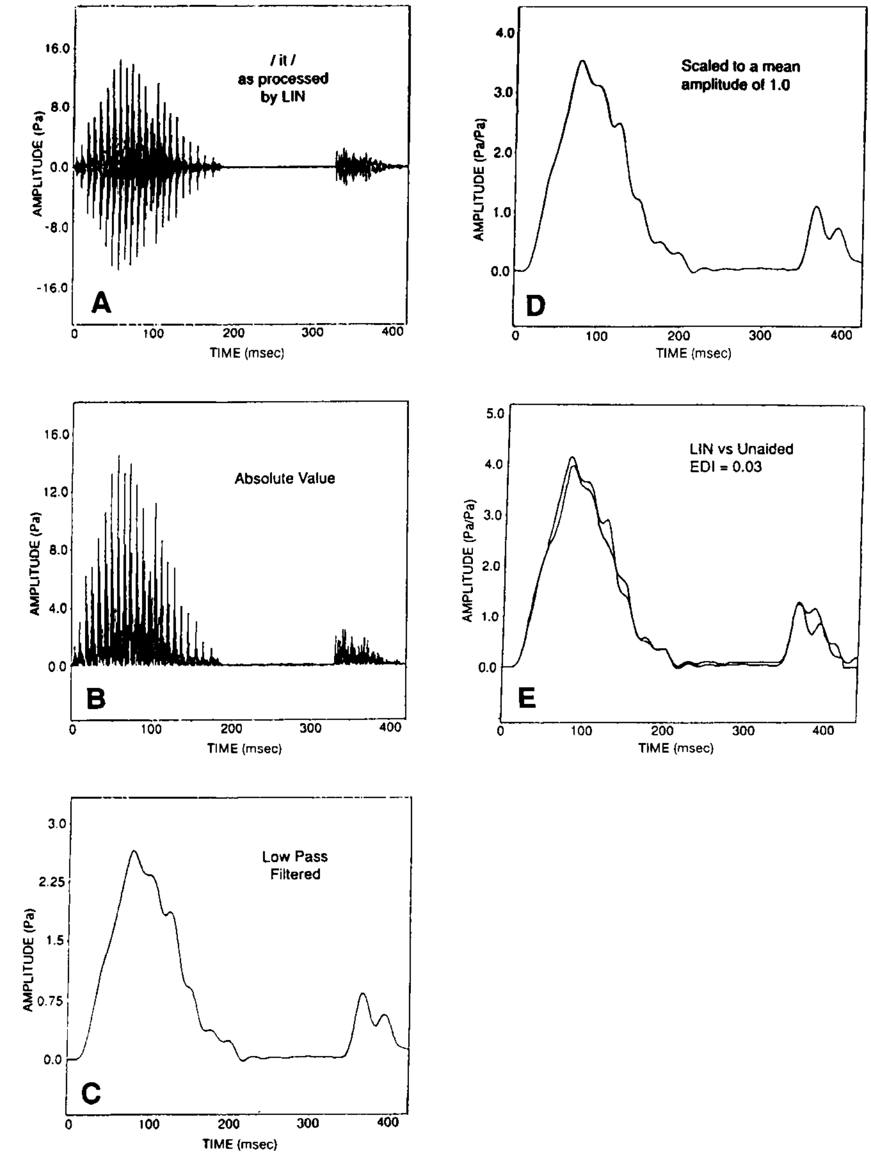

Figure 1. Method of EDI calculation. A) Temporal waveform of the stimulus/it/, as processed by a Class D linear hearing aid and recorded in the ear of one experimental subject. B) Absolute value of the waveform in panel A. C) Temporal envelope of /it/, produced by lowpass filtering the waveform in panel B. D) Temporal envelope of /it/, scaled to a mean amplitude of 1.0. E) Comparison of the aided envelope from D) with the unaided envelope of the same syllable. Additional details regarding EDI calculations are provided in the text.

method. Before this comparison can be made however, the issue of amplitude must be addressed. Since the EDI is based on envelope subtraction, it becomes necessary to scale all envelopes to a common reference point, so that temporal effects may be analyzed without contamination by amplitude variations that usually occur between a single unaided condition, and perhaps numerous aided conditions representing different hearing aids. Scaling is accomplished by calculating the overall mean amplitude of an envelope, and dividing each sampled point of the envelope by the mean amplitude. This process scales each envelope to a mean amplitude of 1.0 (Panel D), allowing a direct comparison of stimulus envelopes. Panel E shows one such comparison, in this case between the aided and unaided envelopes of the syllable /it/. Both waveforms in the panel had been scaled as described, and a technique based on cross correlation was used to verify that they maximally coincided. These steps allow the overall temporal difference that exists between the two waveforms to be calculated.

The EDI is calculated using the following formula:

$$EDI = \left(\sum_{n=1}^{N} |Env1_n - Env2_n| \right) / 2N$$

The formula shows that the unaided envelope (Env2) is subtracted, point by point, from the aided envelope (Env1), and the absolute values of the differences are taken. The EDI is calculated as the mean of these absolute values, divided by 2. Division by 2 places the EDI on a scale that ranges from 1.00

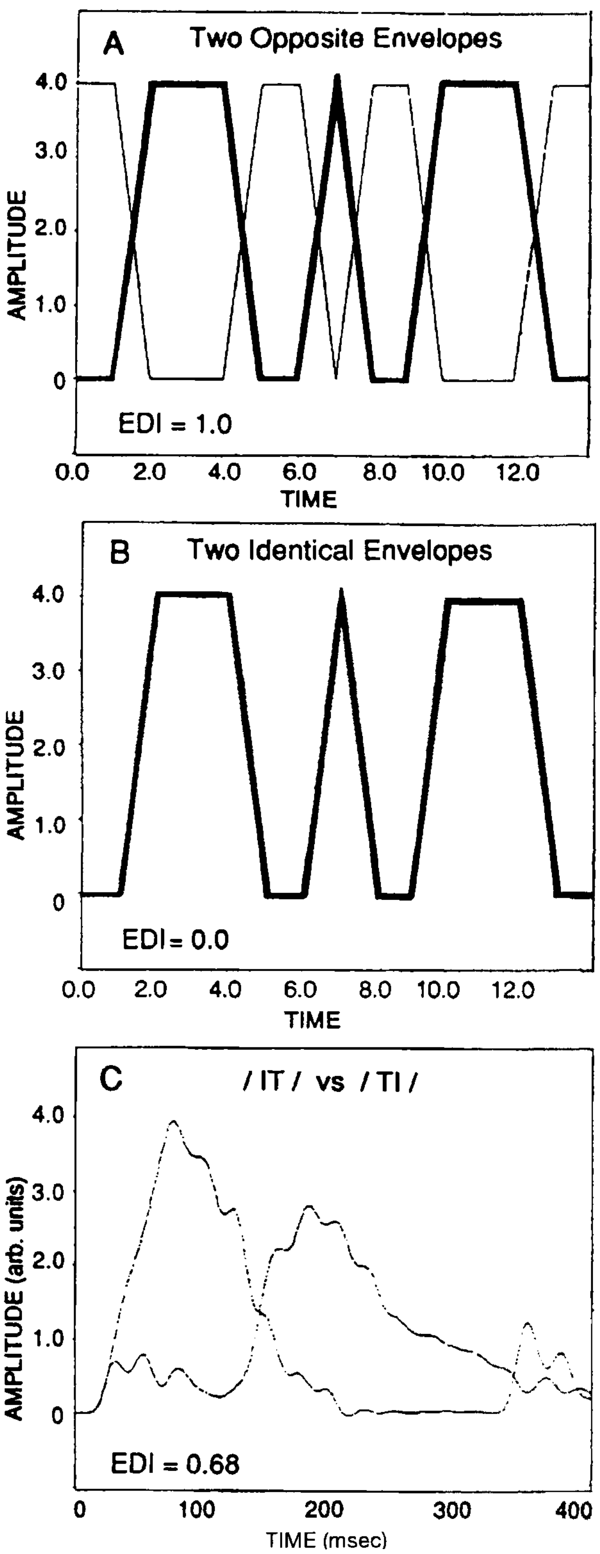

Figure 2. Examples of envelope comparisons and EDI calculations. A) Two mirror-image envelopes, producing an EDI of 1.0. B) Two identical envelopes, superimposed, producing an EDI of 0.00. C) Comparison of the temporal envelopes of the syllables /it/ and /ti/, producing an EDI of 0.68.

(no correspondence between envelopes) to 0.00 (perfect correspondence between envelopes). The EDI for the example shown was 0.03, reflecting excellent agreement between envelopes. EDI calculations such as these may be made for samples of speech, samples of noise, or in general, whenever a precise calculation of the temporal contrasts between two repeatable stimuli is required.

As stated above, the EDI may range from 1.00 to 0.00. Figure 2 provides three examples of envelope comparisons and their resulting EDI values. Panel A shows two contrived envelopes that were deliberately made opposite to one another. Applying the EDI calculations to these envelopes generates an EDI of 1.00. By inverting one of these envelopes (Panel B), the correspondence between envelopes becomes perfect, and the EDI becomes 0.00. Panel C shows the envelopes of /it/ and /ti/, two syllables that would not normally be compared. It is quite clear that very little temporal correspondence exists between these two syllables, and an EDI of 0.68 is produced. By analyzing EDIs, not of dissimilar syllables, but of identical syllables processed by different hearing aids, it becomes possible to determine which circuit algorithms produce the largest or the smallest temporal effects. When this information is compared with behavioral responses that may easily be obtained simultaneously with real-ear recordings, it becomes possible to assess whether the temporal alteration of speech induced by amplification enhances or degrades speech intelligibility.

Example of a Speech Intelligibility Application

Figure 3 shows examples of the temporal alteration of speech and how this could affect speech intelligibility. The left panels of the figure each show the aided and unaided envelopes of the stimulus /shi/, as recorded in the ear of an experimental subject. Panel A represents nonlinear hearing aid signal processing. Panel B represents linear signal processing. Panels C and D show comparable waveforms of the syllable /ith/. Each hearing aid was adjusted to the NAL (Byrne & Dillon, 1986) target frequency response associated with the listener's gradually-sloping high-frequency hearing loss. Each syllable was presented to the listener at 65 dB SPL. EDIs, aided consonant levels (in dB sensation level), and two behavioral responses (test, retest) obtained for each syllable are shown in the table at the bottom of the figure. For the syllable /shi/, the nonlinear algorithm generated an EDI of 0.15. Most of the temporal change (re: unaided) occurred during the vowel, a result that was consistent with the design of the circuit. The aided consonant level of this

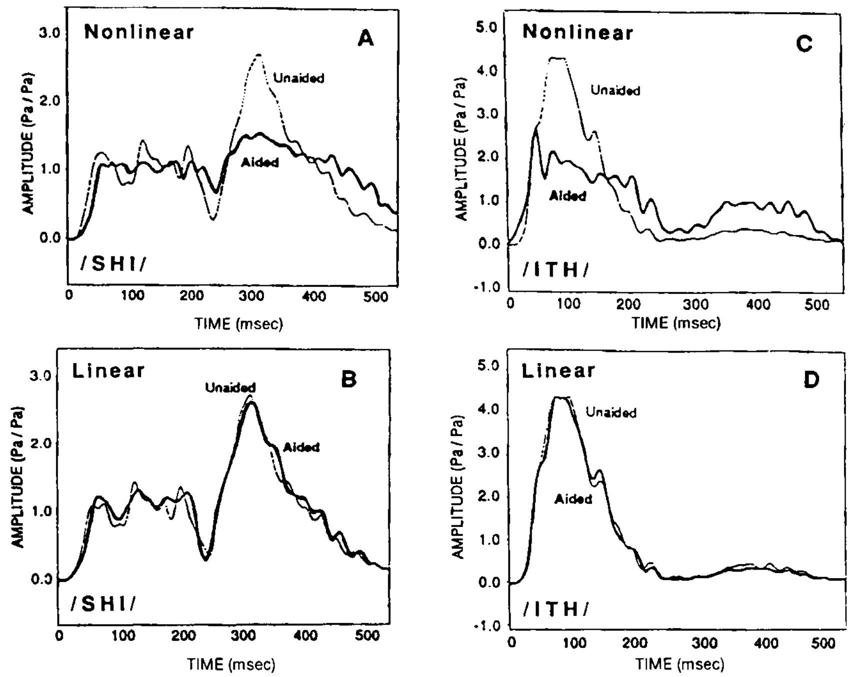

Figure 3. Examples meant to illustrate how the EDI may be applied to the issue of aided speech intelligibility. Each panel shows the aided and unaided envelopes of the indicated syllable, as recorded from the ear of one subject. Nonlinear and linear circuit algorithms are represented in the upper and lower panels, respectively. EDI values, aided consonant levels (in dB SL), and two behavioral responses (test, retest) are included in tabular form below the figure. Additional details are provided in the text.

Syllable	Circuit	EDI	Cons. Level (dB SL)	Response 1	Response 2
/SHI/	Nonlinear	0.15	26	Incorrect	Incorrect
	Linear	0.05	21	Correct	Correct
/ITH/	Nonlinear	0.33	7	Correct	Correct
	Linear	0.04	3	Incorrect	Incorrect

syllable was well above threshold (26 dB SL), but the behavioral responses associated with this algorithm were incorrect on both trials. In contrast, the linear circuit produced a lower consonant level (21 dB SL), and also a lower EDI (0.05). For the linear circuit, however, both behavioral responses were correct. One conclusion that could be drawn from this example is that for this individual, greatest intelligibility occurred when the temporal characteristics of the stimulus were preserved.

In contrast, panels C and D portray an example of possible benefit that the temporal alteration of speech may have on intelligibility. Again comparing EDIs and consonant levels reveals that the nonlinear circuit produced more change in the temporal envelope and a higher consonant level than the linear algorithm. In this instance, the listener responded correctly on both trials when wearing the nonlinear instrument. The linear circuit produced almost no temporal change in the syllable's temporal envelope, but it also failed to raise the consonant to a level much above threshold. For this circuit, both behavioral responses were incorrect. One conclusion that could be drawn from this example is that when consonants are close to threshold, intelligibility may improve if a small increase in audibility is brought about by temporal alteration.

The examples provided in Figure 3 are not intended to resolve the complex questions of how the temporal alteration of speech will affect its intelligibility. Rather, they serve to illustrate how the EDI technique may be used to quantify the temporal changes that in fact occur as a result of amplification. As such, the EDI represents an analysis tool that may be quite powerful in addressing these issues.

Additional EDI Applications

The EDI is not limited to the examination of speech, but may be used to examine the temporal properties of any acoustic signal, acquired under any set of conditions. For example, the EDI may be used to evaluate the extent to which impulsive-type sounds cause hearing aid saturation. An example of this application is shown in Figure 4. Panel A shows the temporal waveform of a hammer striking a nail. This waveform was recorded using a probe microphone that had been placed within a KEMAR's ear canal. The stimulus was presented at 75 dB SPL and the recording was obtained unaided. Panel B shows the envelope of this waveform, derived by methods described earlier. Panel C shows the waveform and Panel D the envelope of the same hammer strike, obtained under conditions of hearing aid satu-

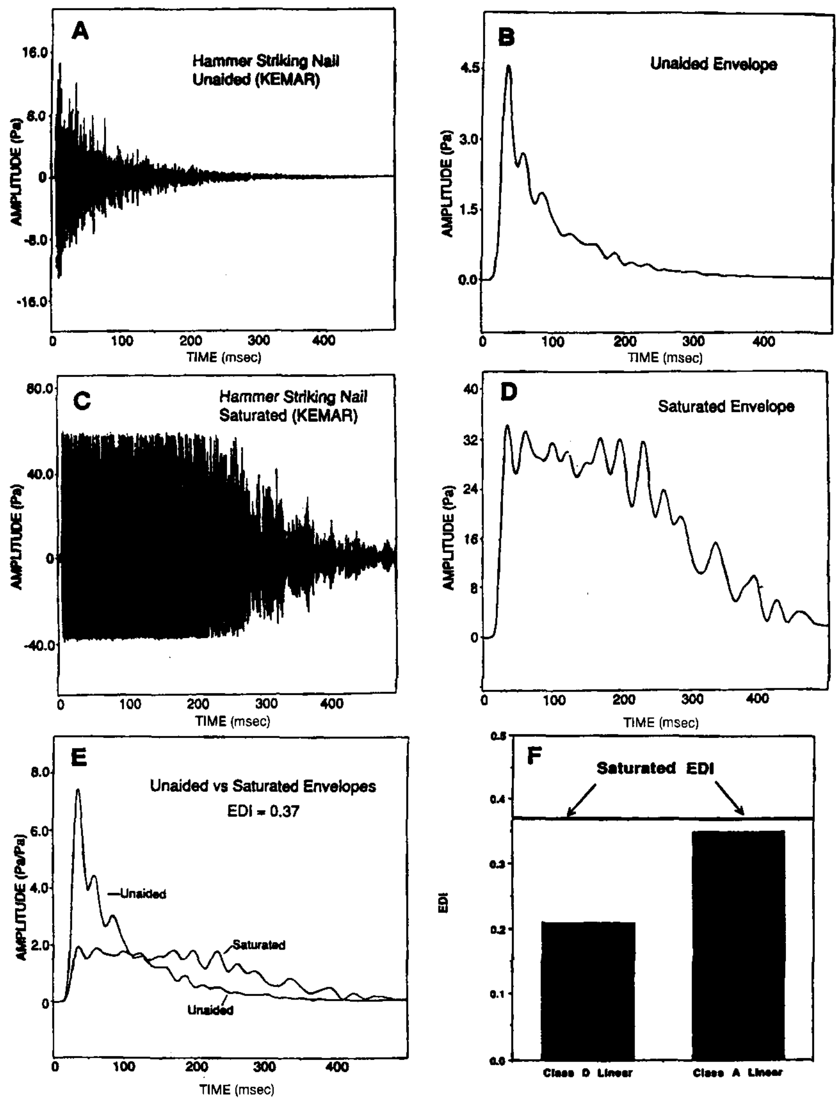

Figure 4. The effect of hearing aid saturation on the EDI. A) Temporal waveform of the sound of a hammer striking a nail, obtained unaided on KEMAR. B) Temporal envelope of the waveform shown in A. C-D) Temporal waveform (C) and envelope (D) of the hammer strike, obtained under conditions of hearing aid saturation. E) Comparison of the unaided and saturated envelopes. F) EDIs associated with a Class D linear and a Class A linear hearing aid, relative to the saturated EDI.

ration. Saturation was produced by adjusting a Class D linear hearing aid (HFA SSPL90 of 118 dB SPL) to produce 40 dB of noise gain (ANSI, 1992), and presenting the stimulus at 95 dB SPL. Panel E compares the unaided and saturated envelopes, after both had been scaled to a mean amplitude of 1.0. In this example, the EDI associated with saturation was 0.37. Similar EDIs have been found for a variety of stimuli, when tested under similar conditions. Panel F shows the real-ear EDI associated with the same Class D linear hearing aid, and also that associated with a Class A linear instrument with a HFA SSPL90 of 104 dB SPL. In this case the stimulus was presented at 70 dB SPL and both hearing aids were adjusted to 30 dB of noise gain. The data suggest that the Class D hearing aid processed the signal with minimal saturation-induced distortion, whereas the same signal may have saturated the Class A instrument. By comparing EDIs with the EDI associated with saturation, it becomes possible to evaluate whether behavioral data obtained under experimental conditions may have been influenced by the presence of saturation-induced distortion.

The EDI may also have applications outside the realm of hearing aids. Figure 5 illustrates a room acoustics application of the EDI. This figure shows the temporal waveform of a transient stimulus (.03 msec pulse) recorded in a reverberant room (Panel A), and in the same room covered with 4-inch acoustic panels (Panel B). The waveform obtained in the reverberant room shows a series of peaks, representing the initial pulse and numerous reflections. The reflections are absent from the waveform obtained in the damped environment. A comparison of the two envelopes (Panel C) produced an EDI of 0.21. Successful damping will result in relatively high EDIs, particularly when undamped waveforms show reflections that are numerous

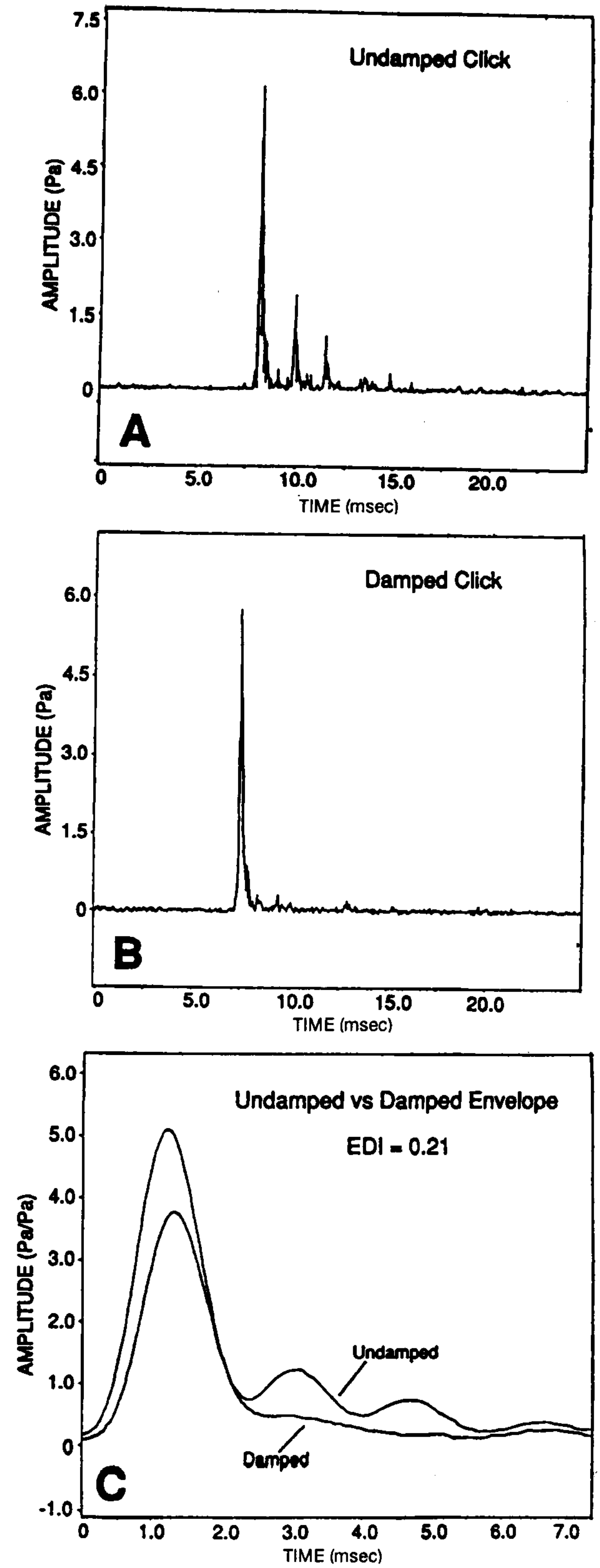

Figure 5. Example of a room acoustics EDI application. A) Temporal waveform of a transient stimulus recorded in a reverberant room. B) Temporal waveform of the stimulus recorded in the same room after acoustic panels had been applied to the walls and ceiling. C) Comparison of the undamped and damped envelopes.

and pronounced. In other words, the EDI will be highest when the untreated room is highly reverberant, and the treated room highly damped. Thus, by monitoring the EDI as damping materials are varied or repositioned, maximum benefit may be achieved.

Summary and Conclusions

The Envelope Difference Index represents a precise quantification of the temporal contrast that exists between two waveforms. The EDI technique provides a way to precisely determine the extent to which different hearing aids alter the natural temporal characteristics of speech. As such, the EDI may be quite useful in determining whether temporal alterations caused by amplification enhance or degrade speech intelligibility. The EDI, however, is not limited to the examination of speech, but may be applied whenever a precise quantification of the temporal contrasts that exist between two temporal waveforms is required.

The EDI technique has several characteristics that make it useful for a variety of applications. First, the EDI is precise; it represents the exact temporal difference between two waveforms. Second, the EDI may be applied to stimuli of any length, from fractions of a phoneme to complete speech passages. Third, the EDI may be used with any acoustic signal, regardless of its temporal characteristics. Finally, the EDI of a test signal may be directly compared with a corresponding behavioral response. Thus, although the utility of the EDI technique is still being assessed, the method seems to hold promise as an effective means of comparing the temporal characteristics between waveforms.

Address correspondence and reprint requests to Dr. Todd Fortune, Argosy Electronics, 10300 W 70th Street, Eden Prairie, MN 55344.

Paper presented at the 2nd International Hearing Aid Conference, Iowa City, June, 1993.

Received October 1, 1993; accepted November 1, 1993.

References

ANSI. Testing hearing aids with a broad-band noise signal. S3–42. New York:American National Standards Institute, 1992.

Byrne D and Dillon H. The National Acoustic Laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear Hear 1986;7:257–265.

Houtgast T and Steeneken HJM. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. J Acoust Soc Am 1985;77:1069–1077.

Moore BCJ. How much do we gain by gain control in hearing aids? Acta Otolaryngol(Suppl) 1990;469:250–256.

Moore BCJ. Characterization and simulation of impaired hearing: Implications for hearing aid design. Ear Hear 1991; 6:154S–161S.

Plomp R. The negative effect of amplitude compression in mul-

tichannel hearing aids in the light of the modulation transfer function. J Acoust Soc Am, 1988;83:2322–2327.
Villchur E. Comments on "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]. J Acoust Soc Am 1989;86:425–427.