Application of the Envelope Difference Index to spectrally-sparse speech

Pamela Souza[1], **Eric Hoover**[1], and **Frederick Gallun**[2]

[1]Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL

[2]National Center for Rehabilitative Auditory Research, Department of Veterans Affairs, Portland, OR

Abstract

Background—Amplitude compression is a common hearing aid processing strategy that can improve speech audibility and loudness comfort but also has the potential to alter important cues carried by the speech envelope. In our previous work, a measure of envelope change, the Envelope Difference Index (EDI; Fortune, Woodruff & Preves, 1994) was moderately related to recognition of spectrally-robust consonants. This follow-up study investigated the relationship between the EDI and recognition of spectrally-sparse consonants.

Method—Stimuli were vowel-consonant-vowel tokens processed to reduce spectral cues. Compression parameters were chosen to achieve a range of EDI values. Recognition was measured for 20 listeners with normal hearing.

Results—Both overall recognition and perception of consonant features were reduced at higher EDI values. Similar effects were noted with noise-vocoded and sine-vocoded processing, and whether or not periodicity cues were available.

Conclusion—The data provide information about the acceptable limits of envelope distortion under constrained conditions. These limits can be applied to consider the impact of envelope distortions in situations where other cues are available to varying extents.

Modern hearing aids employ a variety of signal processing schemes which aim to improve speech recognition for listeners with hearing loss. A hallmark of those schemes is use of wide-dynamic range compression (WDRC). The rapidly varying gain used in fast-acting WDRC systems improves audibility and loudness comfort (see Souza, 2003 for a review), but also modifies acoustic cues. In particular, WDRC affects the envelope[1] by reducing the modulation depth and by introducing overshoot and undershoot caused by the time lag of the compressor. In many respects, these changes are positive; providing more gain to low-intensity phonemes can improve audibility (and therefore intelligibility) even as it reduces modulation depth. However, altering the envelope beyond a certain point may also have a negative effect, and several studies have demonstrated that some envelope alterations can degrade speech recognition (e.g., Van Tasell & Trine, 1996; Souza & Turner, 1996; 1998; Drullman, Festen & Plomp, 1994).

How can we determine when alteration of the envelope is acceptable, and when it is detrimental? With regard to envelope alterations that are caused by WDRC, it is impractical to constrain envelope alteration by recommending specific parameters because all parameters interact with each other. A study that shows better performance with a short

---

[1]Slow (< 50 Hz) variations in speech amplitude (Rosen, 1992)

release time may appear to conflict with a study that shows better performance with a long release time, when the difference is being confounded by different attack times or compression ratios. The issue is further complicated by the finding that patients vary in their reaction to changes to the signal envelope, and that those changes may be more or less important in different environments (Gatehouse, Naylor & Elberling, 2006).

To explore these issues, it would be convenient to use a metric of envelope alteration that captures the combined effect of all processing parameters. Such a metric could be used to confirm that a given set of parameters are unlikely to be problematic, to evaluate new signal processing schemes, or to assess differences among patients in their tolerance to envelope distortion. Several studies (Jenstad & Souza, 2005; 2007; Saade, Zeng, Wygonski, Shannon, Soli, Alwan, 1995; Walaszek, 2008) have explored use of the Envelope Difference Index (EDI) (Fortune et al., 1994) for this purpose. In its conventional form, the EDI is obtained by rectifying and low-pass filtering the broad-band input and output signals to extract the envelope, then calculating the average difference between envelopes. An EDI of 0 indicates no change in envelope due to compression and an EDI of 1 indicates maximum alteration of the envelope.

Jenstad and Souza (2005; 2007) used various compression ratios and release times to create WDRC-amplified signals with a range of EDI values (relative to unprocessed versions of the same tokens). At high EDIs, speech recognition decreased monotonically. For example, an increase in EDI from 0.25 to 0.34 decreased recognition by about 10% for easy speech materials, and about 20% for more difficult (rapidly spoken) speech materials.

Several points raised by Jenstad and Souza require more study. The first concerns the mechanisms that the listener may be using to compensate for envelope distortions. In the Jenstad and Souza data, listeners had mild-to-moderate loss and presumably had access to non-envelope cues including periodicity[2] and fine structure[3]. Listeners might have been able to shift their attention to alternative cues to offset the impact of envelope distortions. With hearing loss which results in broader auditory filters, spectral cues may be less available (Tyler, Hall, Glasberg, Moore, & Patterson, 1984; Souza, Wright, Bor, submitted), and the importance of the envelope may be increased (Boothroyd, Springer, Smith & Schulman, 1988; Davies-Venn, 2010). Spectral cues will be even more limited for cochlear implant wearers, who typically receive only four to eight channels of spectral information (Friesen, Shannon, Baskent & Wang, 2001). It is possible that when spectral cues are less available (i.e., the signal contains less acoustic redundancy), envelope distortions would be problematic at even lower EDIs.

A few studies have varied the amount of spectral information to determine whether availability of spectral cues will offset the effects of envelope distortion, but the findings are not conclusive. Loizou, Dorman and Tu (1999) found that quantizing speech amplitude into a small number of steps (effectively distorting some amplitude information) reduced speech recognition for six-channel speech, but not for 16 channel speech. Although that manipulation did not reduce modulation depth in the same way as compression would have, it does suggest that distortions of the amplitude envelope may be relatively more important in cases where there is restricted spectral information. In a more direct test of compression when spectral information was limited, Zeng and Galvin (1999) found no effect of compression on consonant recognition or feature perception regardless of the number of spectral channels (4, 10, or 20). Because those studies used different speech materials and different methods of amplitude manipulation that were not modeled on WDRC hearing aids,

[2]Amplitude variations at rates between 50 and 500 Hz which provide information about consonant voicing and manner (Rosen, 1992)
[3]Rapid amplitude variations thought to provide cues to consonant place (Rosen, 1992)

it is not clear whether envelope distortion has a greater impact when redundant cues are not available.

A second issue is the tradeoff between envelope alterations which improve audibility (thereby improving speech recognition) and those which distort the envelope (thereby reducing speech recognition). In Jenstad and Souza (2005), larger EDIs sometimes improved phoneme recognition. This was attributed to improved consonant audibility from the fast-acting compression, but audibility was neither strictly controlled nor quantified in that study. In Jenstad and Souza (2007), where audibility was controlled by matching frequency-gain response across EDI conditions, a larger EDI tended to decrease sentence recognition. It is unclear whether the degraded recognition in the 2007 study was due to the sentence test material, which would have contained more prosodic cues and in which the envelope was relatively more important; or whether the effect of EDI is constant, regardless of test material, as long as it does not interact with audibility.

Third, it is of interest to understand whether envelope changes have feature- and/or consonant-specific effects. Jenstad and Souza (2005) found that altering the envelope affected some consonants more than others. This is an empirical test of Rosen's (1992) proposal that envelope would be most important to consonant manner and voicing, and less important to consonant place. If that premise holds true, we should see that manner and voicing change with increasing EDI, but place does not.

The goal of this study was to explore the consequences of envelope change on speech recognition using compression mechanisms representative of wearable hearing aids, for situations where listeners relied on temporal cues. All processing was done using a software simulation of a widely-used hearing aid DSP platform. To focus on loss of spectral detail without the confounding effects of audibility change, we presented signals processed to limit spectral cues to listeners with normal hearing. Such signals reflect a worst-case scenario: if a given amount of envelope distortion has no detrimental effect when spectral information is very limited, it should be at least acceptable (and perhaps negligible) when spectral cues are readily available. Once we understand the acceptable limits of envelope distortion under constrained conditions, those limits can be applied to consider the impact of envelope distortion in situations where other cues are available to varying extents.

Experiment 1

Methods

Participants—Participants were ten adults with normal hearing, aged 21–27 years (mean age 25.1 years). All participants had hearing thresholds of 20 dB HL (re: ANSI, 2004) or better at octave frequencies between .25 and 8 kHz, had no history of speech or hearing disorders, and spoke English as their sole or primary language. All participants but one (author EH) were naive to spectrally-sparse speech.

Stimuli—Test stimuli were a set of 16 vowel-consonant-vowel syllables, each consisting of a consonant /b, d, g, p, t, k, f, θ, s, ʃ, v, ð, z, Z, m, n/ in an /aCa/ context (Turner, Souza, & Forget, 1995).

Each syllable was produced by four talkers (2 male, 2 female) without a carrier phrase for a total of 64 test items. All tokens were digitally recorded at a 44.1 kHz sampling rate with 16 bit resolution. Each syllable was transformed to signal-correlated noise (Schroeder, 1968) as follows. First, the syllable was digitally filtered into four bands. Crossover frequencies were 440, 1130, and 2800 Hz, and the lower-to-upper frequency range across all bands was 176–7168 Hz. Next, the band output was manipulated by randomly multiplying each digital

sample by +1 or −1. This process removed fine-structure information but preserved envelope and periodicity information up to the limits of the listener's ability to detect such cues. Because this processing resulted in generation of energy outside the bandwidth of the filtered signal, the band was refiltered using the original filter settings and amplified with gain appropriate to correct for the power loss of the second filtering. Finally, the filtered segments were digitally mixed. The resulting four-band signal preserved envelope cues but provided only gross cues to spectral shape. Although a one-band signal would have created a pure test of envelope cues, four bands were selected based on pilot data to achieve a range of scores which avoided floor or ceiling effects. At most, such signals have approximate spectral shape and very coarse consonant-vowel transitions (Souza & Rosen, 2009).

EDI calculation—The Envelope Difference Index (EDI) was calculated using locally-developed Matlab code. In each case, the EDI represented a comparison of signal envelopes for a compressed syllable compared to an uncompressed version of the same syllable. First, a syllable was rectified and digitally low-pass filtered using a Butterworth 6th order filter with a 50 Hz cut-off to obtain the syllable envelope. The envelope was downsampled to a sampling frequency of 6000 Hz and the mean amplitude of the syllable was calculated. Each sampled data point of the envelope was scaled to the mean amplitude by dividing every value by the mean. This provided a common reference for comparing the two envelopes. The same steps were conducted for the second signal. The EDI was calculated using the equation below, where Env1 was the uncompressed SCN version of a given syllable, and Env2 was one of the compressed versions of the SCN-processed syllable:

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

and N = number of sample points in the waveforms

Env1$_n$ = the envelope of the compressed waveform

Env2$_n$ = the envelope of the unprocessed waveform

Selection of compression parameters and target EDIs—Because we were interested in the effects of envelope alteration under conditions similar to those which might occur in practice, we first sought to determine what EDIs might occur in clinically-fit, wearable hearing aids. We drew information from two sets of data. The first (Souza, Hoover, Gallun & Brennan, 2010) measured EDI for speech amplified by wearable hearing aids. All recordings were completed using a KEMAR manikin fit with a commercial behind-the-ear hearing aid coupled to a lucite skeleton (unvented) earmold. The WDRC hearing aid had four compression channels plus output limiting and was programmed with the manufacturer's default frequency-gain and compression parameters for a representative mild-to-moderate audiogram. Digital noise reduction and directional microphone response were disabled. Test signals were the 16 vowel-consonant-vowel nonsense syllables described above, presented in quiet at a 0 degree azimuth in soundfield at input levels of 65 dB SPL (representing conversational speech) and 80 dB SPL (representing loud speech). Across all syllables, EDI values ranged from .05 to .27 (mean .12) for the 65 dB SPL input level and from .06 to .36 (mean .14) for the 80 dB SPL input level.

The second data set was from Walaszek (2008). That study also used a commercial behind-the-ear hearing aid, programmed with manufacturer settings for a mild loss, with no earmold venting and with digital noise reduction and directional microphone response disabled. Both fast-acting and slow-acting WDRC were tested. Speech materials were Danish sentences

presented in a background of a single female talker or in ICRA noise. Output signals were recorded in a Bruel and Kjaer ear simulator and processed to separate speech from noise, allowing the EDI calculations to be based on the target speech. Mean EDIs varied from .12 to .22 depending on the compressor speed and type of background noise. To summarize, two independent studies suggest that commercial hearing aids set to default parameters for listeners with mild-to-moderate loss will produce mean EDI values in the neighborhood of . 1-.2, with higher EDIs for some consonants.

Compression parameters (Table 1) were chosen based on the hearing aid recording data described above and on our previous work (Jenstad & Souza, 2005; 2007) to generate output signals with EDIs that varied at least across the .1-.2 range. Note that these parameters were not intended to mimic any single commercial product. Rather, we sought to create EDIs that would be typical of those that might occur in wearable hearing aids. To subject most of the speech dynamic range to compression and to mimic use of a WDRC (low-compression threshold) hearing aid, the compression threshold was set 25 dB below the RMS level of the input signal.

Each four-band syllable was processed using hearing aid simulation software (Gennum Corporation; GennEM v1.0). The GennEM application is a single-channel compressor that operates in the same manner as compression circuits in wearable hearing aids. Although current hearing aids use multichannel compression, such compression can also introduce channel interactions (and spectral distortion) and the goal here was to isolate the effect of (broad-band) envelope distortion. In addition, the rationale for using multichannel compression is to achieve audibility across frequency, due to variations in the listener's dynamic range (Woods, Van Tasell, Rickert, & Trine, 2006). Because the listeners tested here had normal hearing without such variations, multichannel compression was not needed. As a control, performance was also tested with uncompressed (linear) speech.

Figure 1 illustrates the EDI calculation for the test signals. In each panel, the solid line shows the envelope of the compressed syllable /afa/ and the dotted line shows the uncompressed /afa/ and the condition values indicate increasing amounts of compression (higher compression ratio and/or shorter release time) from Table 1. The primary effect of WDRC for this low-intensity voiceless fricative was to increase consonant amplitude relative to vowel amplitude. The overshoot from the compressor is also visible, particularly at the onset of the initial vowel for condition 2. Compression had little effect on the envelope of /afa/ in condition 2, where EDI was 0.07. For /afa/, the EDI reached a maximum of 0.20 for condition 4 in which the highest compression ratio was combined with the shortest release time.

Test procedure—During a test session, the participant was seated in a double-walled sound booth. The digital signals were converted to analog (TDT RP2) and presented at a level of 65 dB SPL via an ER-2 insert earphone to the subject's right ear. All 16 consonants were displayed in Roman orthography on a touch-screen monitor and the participant was asked to select the consonant heard after each trial.

Each participant completed a familiarization block consisting of 20 trials randomly selected from condition 1, as this was the compression condition with the least envelope alteration (i.e., the longest release time and lowest compression ratio). Feedback was provided in the training phase. A test block consisted of 320 trials (16 consonants × 4 speakers × 5 compression conditions) presented in random order, with feedback. Each participant completed two blocks per compression condition. A confusion matrix representing presentations and responses was obtained for each block.

Results

Proportion correct for each EDI condition is shown in Table 2. The values were converted to rationalized arcsine units (Studebaker, 1985) for further analysis. The pattern of decreasing performance at increasing EDIs was confirmed via one-way repeated-measures ANOVA ($F_{4,36}$=15.43, p=.001)[4], with data grouped by mean EDI value (Table 1). There was no difference[5] between EDIs of .00 and .06 (t_9=1.31, p=.223) or .06 and .11 (t_9=.72, p=.489). Performance decreased as EDI increased from .11 to .18. (t_9=2.6, p=.028) and 3 and 4 (t_9=4.0, p=.003).

A secondary goal was to determine whether specific consonant features differed in their susceptibility to envelope change. Following from Rosen (1992), we expected the EDI to be strongly related to perception of consonant voicing and manner. To test this, we performed an information analysis (Wang & Bilger, 1973) for voicing, place, frication, plosiveness, and nasality. All of the features were defined as present or absent for a specific consonant with the exception of place of articulation which was categorized as labial, interdental, or velar. For example, /ava/ was positive for voicing and frication, and had a labial place of articulation. It is necessary to group data into a confusion matrix representing a set of stimulus-response pairs; these were grouped as in Table 2. The analysis was performed on the confusion matrix representing the final score in each condition and for each participant (5 conditions × 10 participants). Only results from the first iteration were used.

Results are shown in Figure 2. Values on the x axis show the mean EDI (see Table 1). Values on the y axis show proportion of relative transmitted information. For each feature a one-way, repeated-measures analysis of variance was used to examine the effect of increasing EDI.

Voicing was reduced across conditions ($F_{4,36}$=13.37, p=.001), with a reduction in voicing cues at the highest EDIs[6] (EDIs .18 and .23). Place scores were low, consistent with previous work indicating that place was poorly transmitted by spectrally-sparse signals (Boothroyd, Mulhearn, Gong, & Ostroff, 1996; Gallun & Souza, 2008; Souza & Rosen, 2009). Place was reduced across conditions ($F_{4,36}$=15.13, p<.005), with a reduction in place cues at the two highest EDIs . Note that the difference between EDIs of .18 and .23 (p=.027) was not significant after correction for multiple tests, but this is likely to have been constrained by a floor effect given the low overall scores.

With regard to manner, frication decreased with increasing EDI ($F_{4,36}$=27.95, p<.005). Plosiveness also decreased with increasing EDIs ($F_{4,36}$=18.95, p=.001); post-hoc analysis indicated that EDI .06 was statistically equivalent to the linear condition; and plosiveness decreased for EDIs from .11 to .23. The pattern for nasality was slightly different. Although nasality decreased as EDI increased ($F_{4,36}$=8.38, p<.005), post-hoc comparisons indicated a significant decrease only for the .23 EDI. The primary cue to nasality is the spectrum of the nasal murmur (Ohde, 1994) which should be less affected by envelope change.

Finally, we considered whether envelope was more important to some consonants than to others, using a stepwise-regression[7] procedure. Three factors were assessed: the predictive value of EDI; the predictive value of consonant; and a predictor based on the interaction between EDI and consonant[8]. EDI was a significant predictor of performance (p=.003).

---

[4]For this and later analyses, if Mauchley's test was significant, Greenhouse-Geisser adjusted values are reported.
[5]When the main effect was significant, the post-hoc comparisons reported are based on means comparisons where the significance level was quantified by the Bonferroni adjustment.
[6]Reported post-hoc comparisons with Bonferroni-qualified significant differences
[7]In this procedure, predictor variables are entered one at a time. Regression coefficients and tests of significance are calculated at each step. If a variable does not contribute significantly to prediction it is eliminated from the model (Gardner, 2001).

Neither consonant (p=.161) nor the interaction between EDI and consonant (p=.512) was a significant predictor. This suggests that the effect of envelope distortion can be considered as occurring broadly across consonants.

Patterns of response by consonant—Figure 3 shows the likelihood of choosing a particular response. Data are expressed as the percent of time a particular consonant was selected within a condition (i.e., within-condition totals sum to 100%). To interpret these data, consider that each presented consonant comprised 6% of the trials. The percent values in Figure 3 are not direct indicators of performance but rather show response bias and how that bias changes as the envelope is altered with compression. For example, for the linear control condition (solid-fill bars), listeners chose /b/, /f/, and /v/ more often than any other consonant (17%, 13% and 12% of responses, respectively). As increasing compression was applied, there was an increase in the proportion of /s, z, ʃ, ʒ/ responses. This change in the distribution of responses may simply be a result of the compression, in a situation where spectral cues are limited: compression increased the amplitude of the medial consonant; consonants tend to have high-frequency spectra; thus increased high-frequency energy was perceived as a fricative. However, it may also be an effect specific to use of SCN for spectral degradation, which has a dense, noisy spectrum qualitatively similar to an intense fricative. It is therefore unclear whether listeners perceived /s, z, ʃ, ʒ/ due to inherent temporal envelope cues, or the use of a noisy signal with a fricative-like quality. In Experiment 2, this question was addressed by using a qualitatively different method of spectral degradation employing a sine wave carrier.

Experiment 2

In order to confirm and extend the results of Experiment 1, a second experiment was performed. In Experiment 1, SCN processing was used to restrict spectral cues. The SCN signals degraded spectral cues and remove fine structure, but also contained periodic variations in amplitude which might have been redundant with envelope cues. Voicing, for example, is weakly cued by envelope but strongly cued by periodicity (Souza & Rosen, 2009). For this reason, conditions were included in Experiment 2 in which the amount of periodicity was systematically varied. This allowed us to explore the extent to which compression effects interact with the presence of periodicity information.

The second goal of Experiment 2 was to examine the extent to which responses in Experiment 1 were influenced by the signal processing used to limit spectral cues. Consider the fricatives /s, z, ʃ, ʒ/ which were chosen infrequently in the linear condition but more often as the effect of the compressor increased (Figure 3). Those consonants are characterized by high-intensity, high-frequency frication noise. However, SCN signals also have a noise-like quality. It is possible that the pattern of results reflects an interaction between compression and the use of SCN processing, rather than effects of compression alone, and such an effect could limit generalization of results. In order to test this systematically, performance was compared when the syllables to be identified were processed using a variety of carrier types.

Methods

Participants—Ten adults with normal hearing, aged 18–42 years (mean age 25 years) were recruited. All participants had hearing thresholds of 20 dB HL or better (re: ANSI, 2004) at octave frequencies between .25 and 8 kHz, had no history of speech or hearing disorders, and spoke English as their sole or primary language. One of the subjects (author

---

[8]The alpha-level criteria for probability of entry into the model was .05 and the probability of removal from the model was .10.

EH) had participated in Experiment 1. The remainder had no experience with vocoded or SCN speech.

Stimuli—Test stimuli were the set of 16 vowel-consonant-vowel syllables used in Experiment 1. Only one female speaker was used. Five test conditions were created: (1) signal-correlated noise, using the stimuli from Experiment 1 (2) noise-vocoded with 30 Hz envelope smoothing filter; (3) noise-vocoded with 300 Hz envelope smoothing filter (4) sine-vocoded with 30 Hz envelope smoothing filter and (5) sine-vocoded with 300 Hz envelope smoothing filter. In vocoded speech, the signal envelope is extracted, filtered, and used to modulate a sine or noise carrier. This process removes fine-structure information and largely preserves envelope; the inclusion of periodicity information depends on the choice of cut-off frequency of the envelope-smoothing filter. It differs from signal-correlated noise in that the rate of amplitude variations can be controlled. In this case, the purpose of the 30- vs. 300-Hz filter was to control the rate of amplitude variations, such that the 300-Hz filter preserved both envelope and periodicity information, and the 30-Hz filter preserved envelope information. The purpose of the noise- vs. sine-vocoded stimuli was to assess contributions of carrier quality, particularly to fricative manner.

The vocoded conditions were created as follows. Each file was digitally filtered into four bands, using sixth-order Butterworth IIR filters. Crossover frequencies were 392, 1005, and 2294 Hz, and the lower-to-upper frequency range across all bands was 100–5000 Hz[9]. Next, the output of each band was half-wave rectified and low-pass filtered (fourth-order Butterworth) at either 30 or 300 Hz to extract the amplitude envelope. The envelope was then multiplied by a carrier, either a tone at the band center frequency, or a noise. The resulting signal (envelope × carrier) was filtered using the same bandpass filter as for the first filtering stage. RMS level was adjusted at the output of the filter to match the original analysis and the signal was summed across bands.

Each syllable in each condition was processed using the same hearing aid simulation software and compression parameters as in Experiment 1. An EDI value was calculated for the compressed syllable relative to its uncompressed counterpart.

Test procedure—Test equipment and procedure were as described for Experiment 1, with the following differences. After 20 familiarization trials with feedback, the subject completed two test blocks, each consisting of 400 trials (16 consonants × 1 speaker × 5 amplification conditions × 5 signal conditions) presented in random order, with feedback. A confusion matrix representing presentations and responses was also obtained for each subject and condition.

Results

Relationship between EDI and performance—Results are shown in Figure 4 and Table 3. The results described here are based on a two-way ANOVA, with repeated measures factors of stimulus condition and compression condition (Figure 4). The analysis was completed on the rationalized-arcsine unit transformed values. The main effects of stimulus condition ($F_{4,36}=13.53$, $p<.005$) and compression condition ($F_{4,36}=19.39$, $p<.005$) and the interaction ($F_{16,144}=1.79$, $p=.038$) were significant. The interaction was explored by completing simple main effects analyses either of the effect of compression within a

---

[9]Filter specifications for the vocoded speech differed slightly from those used in Experiment 1, due to differences in the methodology used to create the stimuli. Previous comparisons of filter settings for vocoded speech suggest this was a negligible difference (Shannon, Zeng, & Wygonski, 1998).

stimulus type, or of the effect of stimulus type within a compression condition. Statistical findings are reported in Table 4, and summarized below.

1. With the exception of the control (linear) condition, all stimulus conditions were subject to distortion of the envelope from compression, and all showed a decrease in performance relative to the control condition particularly when mean EDIs were greater than .11.

2. In general, performance was best for the sine/300 Hz condition and worst for the sine/30 Hz condition. This was consistent with data from Souza and Rosen (2009) and suggested that listeners used the periodicity cues and/or spectral sidebands in the sine/300 Hz condition to improve consonant recognition.

3. There was no significant difference between the two noise carriers and the SCN conditions. This was also consistent with Souza and Rosen (2009) who noted that in the presence of a noise carrier, allowing transmission of periodicity offered minimal benefit over the envelope alone. Presumably the masking properties of the carrier fluctuations minimize the advantage of periodic variations in the envelope for consonant identification (e.g., Whitmal, Poissant, Freyman, Helfer, 2007).

4. For both the 30 Hz cutoff (envelope-only) and the 300 Hz cutoff (envelope + periodicity), there was no significant difference between the noise and sine carrier.

Effect of compression on consonant features—Similar to Experiment 1, we completed an information analysis on the confusion patterns. Because only a single speaker was used, results were collapsed across the 10 subjects resulting in one confusion matrix for each of the 5 EDI conditions (Table 1) × 5 stimulus conditions. Results of the feature analysis are shown in Table 5. The following effects were noted:

1. Voicing was highest for the sine/300 condition and lowest for the noise/30 and sine/30 conditions. The difference was not surprising, because periodic variations in amplitude should be a stronger cue to voicing than envelope alone, and those cues would not have been available with a 30 Hz smoothing filter cutoff (Souza & Rosen, 2009). Voicing also decreased with increasing compression; by an EDI of .23, voicing is similarly poor for all stimulus types except the sine/300 condition, in which periodicity should have been resistant to compression.

2. Consistent with Experiment 1, place was low in all conditions and especially at high EDIs. Transmitted information was worst for the sine/30 condition and highest for the sine/300 condition. Place was similar for the three noise conditions, in which only gross spectral shape would be available (Souza & Rosen, 2009).

3. For uncompressed speech, manner was higher for the conditions which contained envelope plus periodicity cues (sine/300, noise/300, and SCN) than for the 30-Hz vocoded conditions. The differences were reduced at higher EDIs. As noted by Souza and Rosen (2009), combining a sine carrier with a high smoothing filter cutoff has the advantage of providing high-rate amplitude variations as well as spectral sidebands while excluding spurious modulations from a noise carrier.

Patterns of response by consonant—Figure 5 shows response patterns, plotted by consonant. In general, these are similar across processing type, and consistent with the response patterns seen in Experiment 1 (Figure 3). Even for the non-noise (sinusoidal) carrier, listeners continue to choose strong fricative consonants more frequently as the amount of compression increases. This suggests a general effect of compression processing, rather than an artifact due to use of a noise carrier. Balakrishnan, Freyman, Chiang, Nerbonne and Shea (1996) noted a similar effect when consonant-vowel ratio was

manipulated directly, in a manner akin to amplitude compression. Balakrishnan suggested that syllables with naturally-high consonant-vowel ratio in their natural state might benefit from processing that exaggerates that feature, while the same manipulation degrades syllables with naturally low consonant-vowel ratio.

Discussion

Consequences of Reduced Acoustic Redundancy—Consider a continuum in which signals range from high to low redundancy, and where the level of redundancy is controlled by access to mutual information from acoustic and linguistic sources, such as temporal or spectral acoustic cues, or language-dependent lexical or syntactic linguistic cues. In this study, stimuli were processed to limit spectral cues. When fewer channels are available, performance will generally be worse because fewer envelope channels are provided, but there will also be reduced representation of the within-channel temporal envelope. Therefore, the spectral and temporal cues cannot be wholly separated. Nonetheless, we are certainly removing information when vocoding speech or processing it as signal-correlated noise, and in doing so have lowered acoustic redundancy especially in the spectral domain.

How does this apply to listeners with hearing loss? A listener with sensorineural hearing loss would be expected to have more than four channels of spectral information so the processing used mimics spectral degradation that is more representative of a cochlear implant wearer than a listener with sensorineural hearing loss (Friesen et al., 2001). On the other hand, listeners with sensorineural impairment do listen under spectrally degraded conditions (e.g., Leek, Dorman & Summerfield, 1987; Leek & Summers, 1996; Lentz, 2006; Turner & Holte, 1987) so will likely depend to a somewhat greater extent on temporal cues, including envelope. In that sense, the signals used here model temporal cue dependence and loss of spectral information.

Another aspect of redundancy is the content of the speech materials. When high-context materials are used, the listener can apply top-down processing to understand speech even if all acoustic cues are not readily available. However, that process requires greater listening effort and allocation of cognitive resources. This effect was thought to contribute to the findings of Gatehouse et al. (2006) whereby participants who were more negatively affected by envelope distortion (fast release time) tended to be older listeners with poorer working memory. Other situations with low linguistic redundancy listen to speech in an unfamiliar accent or dialect.

The nonsense syllables used here can be considered to have very low redundancy, in that they contained minimal spectral cues and had no lexical or contextual information. The normal-rate sentences used in Jenstad and Souza (2007) were of higher redundancy because they offered spectral and lexical cues. The rapid-rate sentences from Jenstad and Souza would fall in between, with less acoustic redundancy than the normal-rate sentences, but more than the spectrally-limited nonsense syllables.

In the present data, recognition of nonsense syllables with limited spectral cues decreased when EDIs increased from .11 to .18. Jenstad and Souza found that recognition decreased when EDIs increased from .16 to .25 for rapid-rate sentences and from .25 to .34 for normal-rate sentences. Taken together, data across these three studies suggest that envelope distortion is of greater consequence when the signal has less acoustic (spectral) or linguistic redundancy. That idea is also consistent with Cox and Xu's (2010) finding that a short release time (which should produce more envelope distortion) is more problematic when speech has less contextual information (i.e., lower linguistic redundancy) and with data indicating envelope distortion is more detrimental as the spectral content of the signal is

decreased (i.e., lower acoustic redundancy (Fu & Shannon, 1998; Loizou et al., 1999; Drullman et al. 1994a,b; Stone, Fullgrabe & Moore, 2009; Walaszek, 2008).

Audibility versus distortion—When considering the benefit of hearing aid compression for a listener with hearing loss, the envelope distortion represented by the EDI is only one piece of the puzzle. In the work described here, the comparison (uncompressed) signal had the same audibility as the compressed signal because we wished to focus on envelope distortion irrespective of audibility differences. Accordingly, a high EDI was always associated with reduced recognition. In clinical hearing aid fittings, compression will distort the envelope to a degree determined by the specific compression parameters but it will also provide increased audibility of soft speech, particularly low-intensity consonants. In that sense, the EDI value might be viewed as an offset of audibility benefit. When audibility is greatly improved and EDI values are low, the net effect of compression is likely to be positive. When audibility is minimally improved and EDI values are high, the net effect of compression is likely to be negative. When audibility is greatly improved and EDI values are high (as for a listener with severe loss and reduced dynamic range), the net effect of compression is uncertain.

Recent work by Kates (2010) modeled the effects of compression as improved audibility offset by temporal and/or spectral distortion, and proposed an index which incorporates multiple factors. Although Kates' work to date has focused on sentence quality (Arehart, Kates & Anderson, 2010; Arehart, Kates, Anderson, & Harvey, 2007) an index which combines audibility with envelope distortion might also be able to predict speech intelligibility.

Another factor to consider is individual sensitivity to variations in signal amplitude. At high EDIs, recognition is probably reduced because amplitude variations have been diminished, because compression introduces overshoot or undershoot, or both. Subjects with significant loudness recruitment present an interesting case. Recent work by Brennan (2011) demonstrated that when modulated signals are presented through a compression hearing aid, listeners with hearing loss may have the same, better, or poorer modulation sensitivity compared to listeners with normal hearing. The listeners with better modulation sensitivity also had better perception of (WDRC-amplified) speech envelope cues. Although Brennan did not measure loudness growth for his participants, psychoacoustic work suggests that degree of loudness recruitment is related to modulation detection (Moore, Wojtczak, & Vickers, 1996). It is possible that listeners with more recruitment might also respond differently to envelope distortion than listeners with less recruitment.

Consonant feature transmission—Rosen (1992) suggested that the temporal envelope would convey strong cues to consonant manner, weaker cues to consonant voicing (primarily due to the fact that voiced consonants have greater amplitudes than voiceless consonants). Consonant place is expected to be contained in the fine-structure of the signal and not available from envelope information. The addition of periodic information (as in the Sine/300 Hz vocoded stimuli used in Experiment 2) should provide strong voicing cues. Our results support these patterns. In both experiments, place was consistently lower than either voicing or manner. In Experiment 2, voicing improved as the envelope smoothing filter was increased from 30 to 300 Hz. Altering the envelope resulted in a marked decrease in consonant manner and voicing, and a smaller decrease in perception of place. The change in place is interesting, because Rosen's ideas suggest that if place is not conveyed by envelope, altering the envelope should not affect place. However, place can also be conveyed by gross spectral shape (Blumstein, Isaacs, & Mertus, 1982) so if the envelope amplitude is reduced more in one carrier band, gross spectral shape (and therefore place perception) could also change. We suspect this is the reason for the reduced place perception at high EDIs.

Alternative signal representations—How does the EDI compare to other metrics of temporal change? The grandfather of temporal indices, the Speech Transmission Index (STI), is based on the concept that preserving modulations is a desirable outcome. However, as noted by the developers of the EDI (Fortune, et al., 1994), the STI is not appropriate for application to specific elements of speech because its derivation relies on average temporal effects measured over long time intervals. The STI was also devised to capture the effects of noise and/or reverberation in cases where the speech modulations are unchanged, and more recent evaluations of MTF-based indices also noted that they do not capture effects of nonlinear speech processing (Goldsworthy & Greenberg, 2004; Noordhoek and Drullman, 1997).

Gallun and Souza (2008) recently proposed the Spectral Correlation Index (SCI) for use with hearing-aid processed speech. The SCI is a multi-dimensional model that characterizes changes to the signal at different modulation rates and different carrier frequency bands. Although both the EDI and SCI show a relationship to speech recognition (Jenstad & Souza, 2005; 2007; Souza & Gallun, 2010), they represent different theoretical approaches. The EDI characterizes changes to speech in the time domain, and the SCI in the modulation domain. This means that each index will combine specific properties of the envelope in different ways. Our preliminary work (Souza, Gallun, & Hoover, 2009) suggested that the EDI better represents the net effect of interrelated compression parameters, whereas the SCI may be more sensitive to changes in a single compression parameter.

Two multiband implementations of the EDI have also been suggested (Jenstad, Souza & Lister, 2006; Walaszek 2008). Both measured the within-band EDI then averaged across to produce an average EDI. When the average is unweighted, this should produce similar values to the broad-band EDI. For example, Walaszek found a difference of about .03, on average, between the broad-band EDI and an 8-band averaged EDI.

Can the EDI advise clinical practice?—The study described here tested a specific experimental question: when subjects rely on envelope cues, how much can that envelope be altered before recognition is affected? Clinical audiologists seek answers to a similar question when adjusting WDRC hearing aids (Jenstad, Van Tasell & Ewert, 2003). The present data represents a step towards that goal by providing a measure in which we can consider the net effect of parameters on the envelope. Jenstad and Souza (2007) argued that equivalent EDIs derived with any combination of compression parameters should result in similar recognition scores. This contention is supported by the data of Jenstad and Souza (2007) in combination with Walaszek (2008). As one example, both studies processed sentences to create EDI values of .16, but with different compression parameters. Jenstad and Souza combined a higher compression ratio (4:1) with longer release time (800 ms), while Walaszek combined a lower compression ratio (2.5:1) with shorter release time (80 ms). The resulting sentences were recognized correctly 65–70% of the time by Jenstad and Souza's participants, and 65% of the time by Walszek's participants.

The EDI will be most meaningful when considered as a general descriptor or categorization of envelope distortion. To put this another way, the EDI is not so precise that a particular syllable with an EDI of 0.22 will be more difficult to recognize than the same syllable with an EDI of 0.20. Nonetheless, the EDI can represent the overall distortion such that each set of hearing-aid parameters will have an EDI and resulting effect on recognition. A correlate might be the use of the Speech Intelligibility Index (SII) (ANSI, 1997) in hearing aid fitting, where the SII is not applied precisely on a word-by-word, or syllable-by-syllable basis, but is used to categorize and contrast the effect of different hearing aid settings on speech. The SII can then be considered in the context of other factors, such as loudness comfort and perceived speech quality, when making fitting decisions. The EDI might be used in a similar

way. That is, it provides one piece of information in the fitting process. The clinician might then consider that envelope distortion in the context of other fitting decisions: whether the compression improves audibility; whether moderating the compression parameters would cause unacceptably low audibility, or unacceptably high loudness comfort; whether that patient is likely to have ready access to spectral cues/signal redundancy which will offset the envelope distortion.

In summary, the EDI offers a simple and convenient way to capture temporal distortion. With regard to WDRC, the data here and in our previous papers suggest that the optimum set of compression characteristics should not be considered in terms of a specific release time or compression ratio, but as a "balance point" where the negative effect of temporal distortion does not outweigh the positive effect of improved audibility for a particular listener and situation. Although it is not clear whether the EDI can be adapted into a useful clinical tool, it offers one approach to characterize temporal envelope change while studying various other effects. For example, when considered in the context of earlier work, data suggest that the consequences of envelope distortion may be increased in cases of lower acoustic redundancy due to either listener or situational factors. Accordingly, it offers an experimental tool suited to the study of hearing loss and/or device effects.

Acknowledgments

References

ANSI. American National Standards Institute Specification for Audiometers (S3.6-2004). New York: ANSI; 2004.

ANSI. American National Standards Institute Method for Calculation of the Speech Intelligibility Index (ANSI S3.5-1997). New York: ANSI; 1997.

ANSI. American National Standards Institute Specification of Hearing Aid Characteristics (S3.22-2009). New York: ANSI; 2009.

Arehart KH, Kates JM, Anderson MC, Harvey LO Jr. Effects of noise and distortion on speech quality judgments in normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America. 2007; 122:1150–64. [PubMed: 17672661]

Arehart KH, Kates JM, Anderson MC. Effects of noise, nonlinear processing, and linear filtering on perceived speech quality. Ear and Hearing. 2010; 31:420–36. [PubMed: 20440116]

Balakrishnan U, Freyman RL, Chiang YC, Nerbonne GP, Shea KJ. Consonant recognition for spectrally degraded speech as a function of consonant-vowel intensity ratio. Journal of the Acoustical Society of America. 1996; 99:3758–3769. [PubMed: 8655807]

Blumstein SE, Isaacs E, Mertus J. The role of the gross spectral shape as a perceptual cue to place articulation in initial stop consonants. Journal of the Acoustical Society of America. 1982; 72:43–50. [PubMed: 7108042]

Boothroyd A, Springer N, Smith L, Schulman J. Amplitude compression and profound hearing loss. Journal of Speech and Hearing Research. 1988; 31:362–76. [PubMed: 3172753]

Boothroyd A, Mulhearn B, Gong J, Ostroff J. Effects of spectral smearing on phoneme and word recognition. Journal of the Acoustical Society of America. 1996; 100:1807–1818. [PubMed: 8817914]

Brennan, M. Doctoral dissertation. University of Washington; Seattle, WA: 2011. Role of aided temporal cues in aided speech recognition.

Cox RM, Xu J. Short and long compression release times: speech understanding, real-world preferences and association with cognitive ability. Journal of the American Academy of Audiology. 2010; 21:121–38. [PubMed: 20166313]

Davies-Venn, E. Doctoral dissertation. University of Washington; Seattle, WA: 2010. Spectral resolution, severe loss and hearing aids.

Drullman R, Festen JM, Plomp R. Effect of reducing slow temporal modulations on speech reception. Journal of the Acoustical Society of America. 1994a; 95:2670–80. [PubMed: 8207140]

Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech recognition. Journal of the Acoustical Society of America. 1994b; 95:1053–64. [PubMed: 8132899]

Fortune TW, Woodruff BD, Preves DA. A new technique for quantifying temporal envelope contrasts. Ear and Hearing. 1994; 15:93–99. [PubMed: 8194684]

Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. Journal of the Acoustical Society of America. 2001; 110:1150–1163. [PubMed: 11519582]

Fu QJ, Shannon RV. Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. Journal of the Acoustical Society of America. 1998; 104:2570–7. [PubMed: 9821336]

Gallun F, Souza P. Exploring the role of the modulation spectrum in phoneme recognition. Ear and Hearing. 2008; 29:800–813. [PubMed: 18596640]

Gardner, RC. Psychological statistics using SPSS for Windows. New Jersey: Prentice Hall; 2001.

Gatehouse S, Naylor G, Elberling C. Linear and nonlinear hearing aid fittings—2. Patterns of candidature. International Journal of Audiology. 2006; 45:153–71. [PubMed: 16579491]

Goldsworthy R, Greenberg J. Analysis of speech-based speech transmission index methods with implications for nonlinear operations. Journal of the Acoustical Society of America. 2004; 116:3679–3689. [PubMed: 15658718]

Jenstad LM, Van Tasell DJ, Ewert C. Hearing aid troubleshooting based on patients' descriptions. Journal of the American Academy of Audiology. 2003; 14:347–60. [PubMed: 14620609]

Jenstad LM, Souza PE. Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. Journal of Speech, Language and Hearing Research. 2005; 48(3):651–67.

Jenstad LM, Souza PE. Temporal envelope changes of compression and speech rate: combined effects on recognition for older adults. Journal of Speech Language and Hearing Research. 2007; 50(5): 1123–38.

Jenstad, LM.; Souza, PE.; Lister, AM. Development of a metric for quantifying the temporal envelope of speech. Paper presented at the International Hearing Aid Research Conference; Tahoe City, CA. 2006.

Kates JM. Understanding compression: Modeling the effects of dynamic-range compression in hearing aids. International Journal of Audiology. 2010; 49:395–409. [PubMed: 20225931]

Leek MR, Dorman MF, Summerfield Q. Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America. 1987; 81:148–54. [PubMed: 3819173]

Leek MR, Summers V. Reduced frequency selectivity and the preservation of spectral contrast in noise. Journal of the Acoustical Society of America. 1996; 100:1796–1806. [PubMed: 8817905]

Lentz JJ. Spectral-peak selection in spectral-shape discrimination by normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America. 2006; 120:945–56. [PubMed: 16938982]

Loizou PC, Dorman M, Tu Z. On the number of channels needed to understand speech. Journal of the Acoustical Society of America. 1999; 106:2097–103. [PubMed: 10530032]

Moore B, Wojtczak M, Vickers D. Effect of loudness recruitment on the perception of amplitude modulation. Journal of the Acoustical Society of America. 1996; 100:481–9.

Noordhoek IM, Drullman R. Effect of reducing temporal intensity modulations on sentence intelligibility. Journal of the Acoustical Society of America. 1997; 101:498–502. [PubMed: 9000740]

Ohde RN. The development of the perception of cues to the [m]-[n] distinction in CV syllables. Journal of the Acoustical Society of America. 1994; 96:675–686. [PubMed: 7930067]

Rosen S. Temporal information in speech: Acoustic, auditory and linguistic aspects. ICASSP 1997. 1992; 336:367–373.

Saade J, Zeng F-G, Wygonski JJ, Shannon RV, Soli S, Alwan A. Quantitative measures of envelope cues in speech recognition. Journal of the Acoustical Society of America. 1995; 97:3244.

Schroeder MR. Reference signal for signal quality studies. Journal of the Acoustical Society of America. 1968; 44:1735–1736.

Shannon RV, Zeng FG, Wygonski J. Speech recognition with altered spectral distribution of envelope cues. Journal of the Acoustical Society of America. 1998; 104:2467–2476. [PubMed: 10491708]

Souza P. Effects of compression on speech acoustics, intelligibility and speech quality. Trends in Amplification. 2003; 6:131–165.

Souza P, Rosen S. Effects of envelope bandwidth on the intelligibility of sine- and noise-vocoded speech. Journal of the Acoustical Society of America. 2009; 126:792–805. [PubMed: 19640044]

Souza, P.; Gallun, F.; Hoover, E. A comparison of modulation indices for describing amplitude-compressed speech. Paper presented at the Acoustical Society of America; 2009.

Souza P, Gallun F. Hearing aid amplification and consonant modulation spectra. Ear and Hearing. 2010; 31:268–76. [PubMed: 20071993]

Souza, P.; Hoover, E.; Gallun, F.; Brennan, M. Assessing envelope distortion in clinically-fit hearing aids. Paper presented at the International Hearing Aid Conference; 2010.

Souza PE, Turner CW. Effect of single-channel compression on temporal speech information. Journal of Speech and Hearing Research. 1996; 39:901–11. [PubMed: 8898245]

Souza PE, Turner CW. Multichannel compression, temporal cues and audibility. Journal of Speech, Language and Hearing Research. 1998; 41:315–26.

Souza, P.; Wright, R.; Bor, S. Consequences of broad auditory filters for identification of multichannel-compressed vowels. (submitted)

Stone MA, Fullgrabe C, Moore BC. High-rate envelope information in many channels provides resistance to reduction of speech intelligibility produced by multi-channel fast-acting compression. Journal of the Acoustical Society of America. 2009; 126:2155–8. [PubMed: 19894794]

Studebaker GA. A "rationalized" arcsine transform. Journal of Speech and Hearing Research. 1985; 28:455–62. [PubMed: 4046587]

Turner CW, Holte LA. Discrimination of spectral-peak amplitude by normal and hearing-impaired subjects. Journal of the Acoustical Society of America. 1987; 81:445–51. [PubMed: 3558961]

Turner CW, Chi SL, Flock S. Limiting spectral resolution in speech for listeners with sensorineural hearing loss. Journal of Speech, Language and Hearing Research. 1999; 42:773–784.

Turner CW, Souza PE, Forget LN. Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. Journal of the Acoustical Society of America. 1995; 97:2568–76. [PubMed: 7714274]

Tyler RS, Hall JW, Glasberg BR, Moore BC, Patterson RD. Auditory filter asymmetry in the hearing impaired. Journal of the Acoustical Society of America. 1984; 76:1363–8. [PubMed: 6512098]

Van Tasell DJ, Trine TD. Effects of single-band syllabic amplitude compression on temporal speech information in nonsense syllables and in sentences. Journal of Speech and Hearing Research. 1996; 39:912–22. [PubMed: 8898246]

Walaszek, J. Unpublished master's thesis. Denmark: Technical University of Denmark; 2008. Effects of compression in hearing aids on the envelope of the speech signal, Signal based measures of the side-effects of the compression and their relation to speech intelligibility.

Wang MD, Bilger RC. Consonant confusions in noise: A study of perceptual features. Journal of the Acoustical Society of America. 1973; 54:1248–1266. [PubMed: 4765809]

Whitmal NA, Poissant SF, Freyman RL, Helfer KS. Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience. Journal of the Acoustical Society of America. 2007; 122:2376–88. [PubMed: 17902872]

Woods WS, Van Tasell DJ, Rickert MA, Trine TD. SII and fit-to-target analysis of compression system performance as a function of number of compression channels. International Journal of Audiology. 2006; 45:630–644. [PubMed: 17118906]

Zeng FG, Galvin J. Amplitude mapping and phoneme recognition in cochlear implant listeners. Ear and Hearing. 1999; 20:60–74. [PubMed: 10037066]
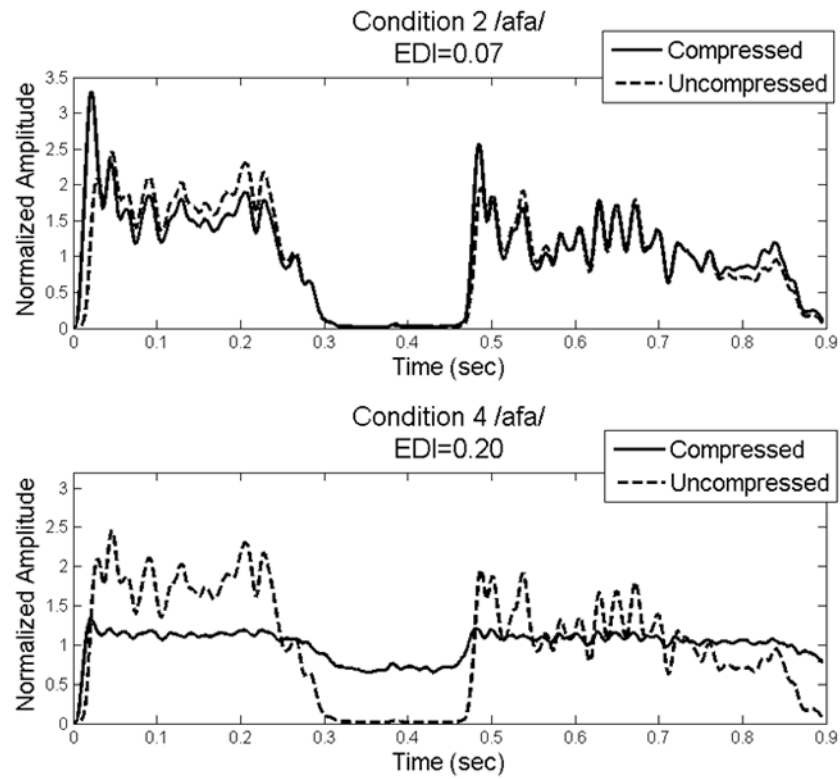
Figure 1.
Example of EDI for compression conditions 2 (top panel) and 4 (bottom panel). In each panel, the solid line shows the envelope for the compressed /afa/ and the dashed line for the uncompressed /afa/.
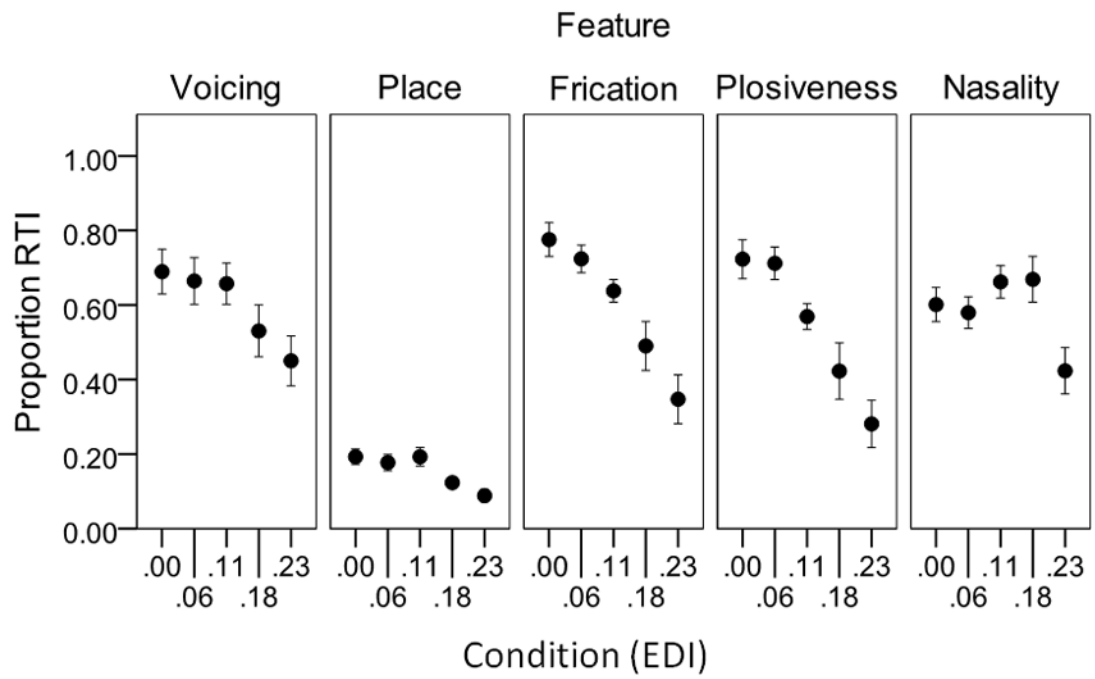
Figure 2.
Proportion of relative transmitted information (RTI) as a function of condition (EDI). Each panel shows results for a different consonant feature. Error bars represent +/1 one standard error about the mean.

Figure 3.
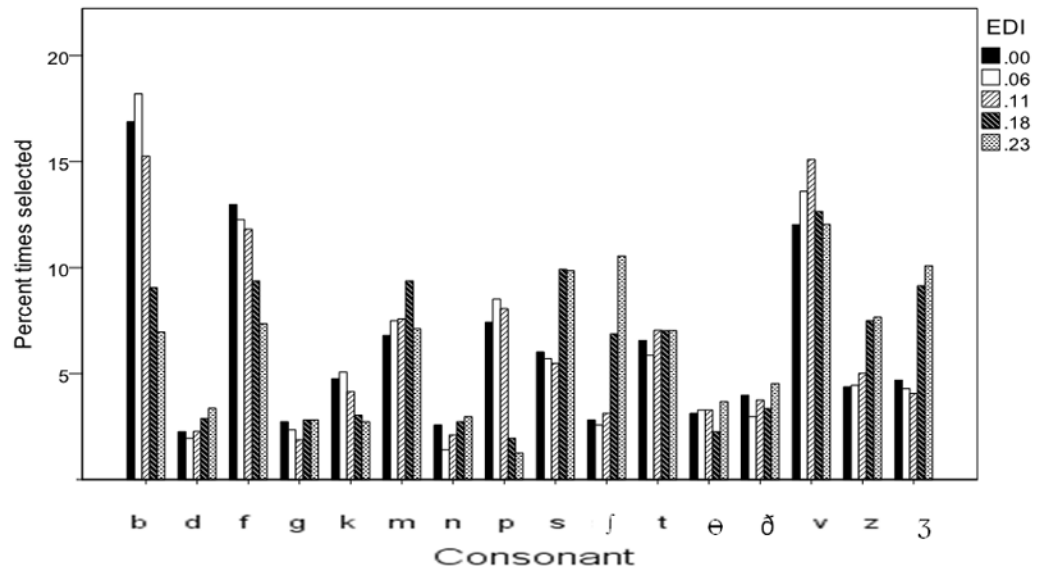Response patterns by consonant. Each EDI condition is shown by a different shading pattern. Values on the y axis represent the percent of times a particular consonant is selected within an EDI condition. For example, in the linear control condition (EDI .00), /b/ was chosen 16.9% of the time, /d/ was chosen 2.3% of the time, and so on; and all of the values for the linear control condition will sum to 100%.
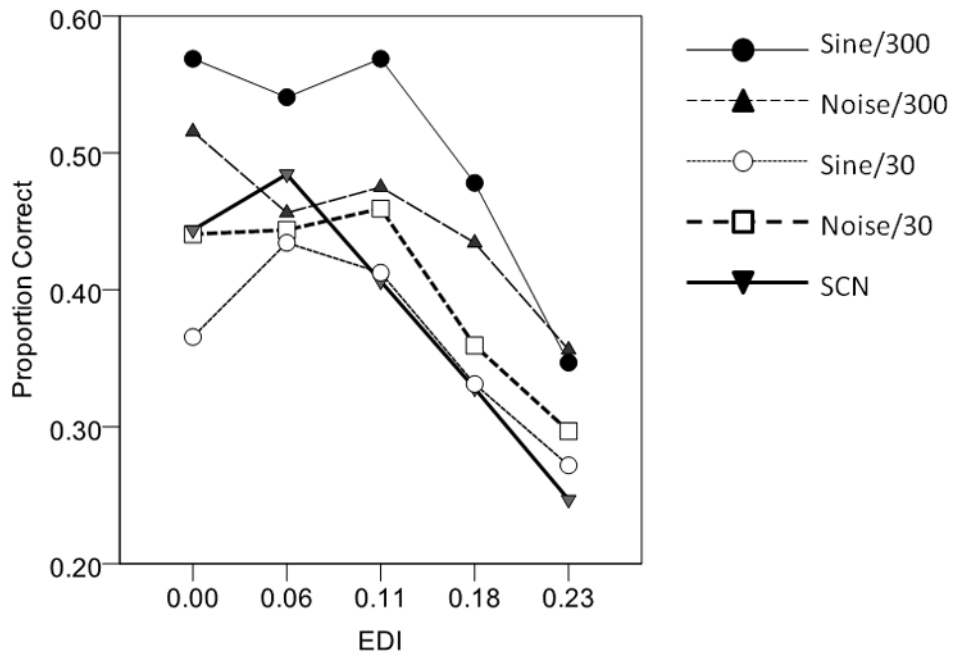
Figure 4.
Proportion correct as a function of condition (EDI). Each symbol/line combination shows results for a different spectrally-limiting processing. For ease of viewing, EDI conditions are evenly spaced on the x axis and variability is provided in Table 4.
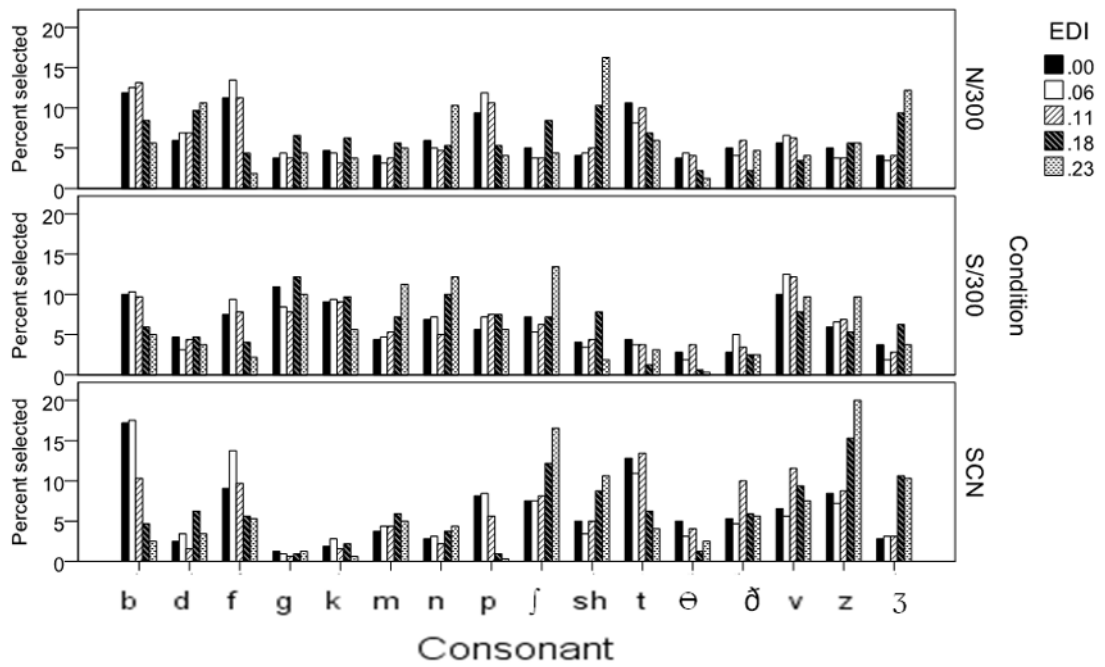
Figure 5.
As in Figure 3, but showing results for three of the signal processing manipulations in Experiment 2: noise carrier, with 300 Hz envelope filter cutoff (N/300); SCN processing (SCN); sine carrier, with 300 Hz envelope filter cutoff (S/300).

Table 1

Compression parameters for the four amplification conditions. In each case, attack time was 6 ms and compression threshold was 40 dB SPL.

	Compression Ratio	Release Time (ms)	Mean EDI
Condition 0	1:1	n/a	.00
Condition 1	2:1	800	.06
Condition 2	4:1	800	.11
Condition 3	4:1	12	.18
Condition 4	10:1	12	.23

Table 2

Mean proportion correct for each amplification condition in Experiment 1.

Condition	Mean EDI	Mean proportion correct	Standard error of proportion correct
0	.00	.45	.04
1	.06	.44	.04
2	.11	.44	.04
3	.18	.38	.03
4	.23	.32	.03

Table 3

Means (and standard deviations) for each signal type and EDI condition in Experiment 2 (refer to Figure 4 for graphical comparison of the data). The values in each cell are the mean proportion correct (standard deviation).

Mean EDI	.00	.06	.11	.18	.23
Sine/300	.57 (.11)	.54 (.10)	.57 (.16)	.48 (.12)	.35 (.12)
Noise/300	.52 (.12)	.46 (.14)	.48 (.09)	.43 (.12)	.36 (.10)
Sine/30	.37 (.12)	.43 (.14)	.41 (.11)	.33 (.09)	.27 (.12)
Noise/30	.44 (.12)	.44 (.17)	.46 (.14)	.36 (.15)	.30 (.12)
SCN	.44 (.11)	.48 (.12)	.41 (.07)	.33 (.12)	.23 (.12)

Table 4

Simple main effects and post-hoc results for Experiment 2

	Simple main effect of compression	$F_{4,49}$	p	EDI conditions significantly different from baseline (compression 0) (Fisher's LSD)
	SCN	6.76	<.005	0.18, 0.23
	Noise/300 Hz	2.45	.060	
	Noise/30 Hz	2.43	.061	
	Sine/300 Hz	5.64	.001	0.23
Condition	Sine/30 Hz	2.97	.029	None. However, EDIs of .06 and .23 were significantly different at p=.004, and EDIs of .11 and .23 were significantly different at p=.010.
	Simple main effect of stimulus type	$F_{4,49}$	p	Stimulus conditions significantly different from each other (Fisher's LSD)
	0	4.23	.005	Sine/300 Sine/30
	1	0.97	.432	
	2	2.93	.031	Sine/300 Sine/30
	3	2.93	.031	Sine/300 Sine/30
Compression	4	1.64	.180	

NIH-PA Author Manuscript   NIH-PA Author Manuscript   NIH-PA Author Manuscript

Table 5

Results of feature analysis for Experiment 2. All values are percent relative transmitted information (maximum possible = 1.0).

Condition		Mean EDI					
		0.00	**0.06**	**0.11**	**0.18**	**0.23**
Noise/300 Hz	Voicing	0.593	0.452	0.575	0.507	0.355
	Place	0.318	0.279	0.293	0.218	0.147
	Manner	0.652	0.586	0.612	0.468	0.41
Noise/30 Hz	Voicing	0.415	0.451	0.421	0.388	0.31
	Place	0.295	0.288	0.348	0.163	0.138
	Manner	0.562	0.559	0.582	0.381	0.325
SCN	Voicing	0.512	0.582	0.583	0.396	0.368
	Place	0.305	0.328	0.253	0.146	0.101
	Manner	0.695	0.6	0.458	0.458	0.36
Sine/300 Hz	Voicing	0.664	0.736	0.711	0.605	0.489
	Place	0.362	0.359	0.373	0.265	0.17
	Manner	0.731	0.732	0.715	0.566	0.451
Sine/30 Hz	Voicing	0.487	0.557	0.498	0.488	0.325
	Place	0.176	0.212	0.207	0.127	0.131
	Manner	0.485	0.526	0.595	0.403	0.396