Acoustical correlates of performance on a dynamic range compression discrimination task

Andrew T. Sabin^{a)}

Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, Illinois 60201

Frederick J. Gallun^{b)}

VA RR&D National Center for Rehabilitative Auditory Research (NCRAR), Portland VA Medical Center, 3710 SW US Veterans Hospital Road, Portland, Oregon 97239

Pamela E. Souza

Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, Illinois 60201

(Received 13 September 2012; revised 25 June 2013; accepted 27 June 2013)

Dynamic range compression is widely used to reduce the difference between the most and least intense portions of a signal. Such compression distorts the shape of the amplitude envelope of a signal, but it is unclear to what extent such distortions are actually perceivable by listeners. Here, the ability to distinguish between compressed and uncompressed versions of a noise vocoded sentence was initially measured in listeners with normal hearing while varying the threshold, ratio, attack, and release parameters. This narrow condition was selected in order to characterize perception under the most favorable listening conditions. The average behavioral sensitivity to compression was highly correlated to several acoustical indices of modulation depth. In particular, performance was highly correlated to the Euclidean distance between the modulation spectra of the uncompressed and compressed signals. Suggesting that this relationship is not restricted to the initial test conditions, the correlation remained largely unchanged both (1) when listeners with normal hearing were tested using a time-compressed version of the original signal, and (2) when listeners with impaired hearing were tested using the original signal. If this relationship generalizes to more ecologically valid conditions, it will provide a straightforward method for predicting the detectability of compression-induced distortions. © 2013 Acoustical Society of America.

[<http://dx.doi.org/10.1121/1.4816410>]

PACS number(s): 43.66.Mk, 43.66.Ba, 43.66.Ts [TD]

Pages: 2136–2147

I. INTRODUCTION

In many situations it is desirable to manipulate the dynamic range of a signal—the differences between the highest and lowest observed intensities. When the goal is to reduce this range, a signal processing technique known as amplitude compression (referred to hereafter as compression) can be used. Compression is commonly used in hearing aids in order to decrease the dynamic range of a signal. Such processing is desirable because it can compensate for the reduced difference between the detection threshold and the uncomfortable listening level that is characteristic of sensorineural hearing loss (for a review, see [Hickson, 1994](#); [Souza, 2002](#)). Compression is also widely used in media production for a variety of goals such as reducing temporal variations in level, increasing loudness, or distorting the temporal envelope in an esthetically desirable way. Whenever compression is applied, some distortion to the temporal envelope is

inevitable (e.g., [Stone and Moore, 1992](#); [Jenstad and Souza, 2005](#); [Stone and Moore, 2007](#)). At present, there have only been a few investigations that attempt to quantify how sensitive listeners are to those distortions, and we are not aware of any that have tried to model this sensitivity (though see [Kates and Arehart, 2010](#) for a model of sound quality that accounts for compression). Such a model could be used to determine which combinations of compressor parameter values (e.g., threshold, ratio, attack, and release values) are perceptually distinct from each other, so that it is not necessary to consider all possible combinations. Here we attempt to quantify behavioral sensitivity to compression-induced distortions to the temporal envelope, and we examine which acoustical indices are correlated to this sensitivity.

The majority of prior work examining the influence of dynamic range compression on perception has focused on how hearing aid compression affects speech intelligibility and quality. These experiments often examine how the use of various compression parameter values affect individuals with hearing loss. Overall, hearing aid compression, in comparison to linear amplification, usually leads to a modest improvement in speech intelligibility for low input level sounds, and less loudness discomfort for high input sounds (for review see [Souza, 2002](#)). Beyond these general effects,

^{a)} Author to whom correspondence should be addressed. Electronic-mail: a-sabin@northwestern.edu

^{b)} Also at: Dept. of Otolaryngology/Head&Neck Surgery Oregon Health & Science University 3181 SW Sam Jackson Park Road, PV01, Portland, OR 97239.

investigations of the influences of specific compression parameter values on perception have been less conclusive. Gatehouse (2006) reviewed thirteen studies, each of which examined the relative benefits of fast vs slow time constants and found several studies supporting each possible result. These seemingly contradictory results could be due to differences in evaluation criteria (Moore, 2008), compression algorithms (Stone *et al.*, 1999), or listening environment (Gatehouse *et al.*, 1999, 2003), as well as individual differences in cognitive ability across participant cohorts (Gatehouse *et al.*, 2003).

In contrast to studies of speech intelligibility and quality, there have been fewer examinations as to how sensitive listeners are to changes in compression parameter values. One of the initial investigations into this topic (Nabelek, 1984) focused on the abilities of listeners to distinguish between uncompressed and compressed sentences, or between various compression parameter values. In all cases, a single channel of compression was used, a single sentence was used, and the loudness was held constant across all sounds. For each tested compressor parameter (attack time, release time, ratio, and input level), there was a monotonic relationship between the parameter value and discrimination performance. Similarly, in a more recent study, investigators examined the discrimination of release time constants using multiple sentences and multiband compression (Gilbert *et al.*, 2008). In this study there was also a monotonic relationship between the manipulated parameter value and the discrimination score. In both investigations the authors noted a large amount of individual-to-individual variation, in some cases ranging from chance to perfect performance for the same condition. Individual differences in training were cited as one source of this variation (Nabelek, 1984). This individual variability prevents general conclusions about sensitivity to compression parameters.

Here we attempted to characterize the upper limit of perceptual sensitivity to compression-induced distortions by using a paradigm with highly favorable listening conditions. Our rationale was that this approach would be likely to reduce individual variability and would yield the upper limit of what is perceivable, thereby providing a principled way to identify how distinguishable compression parameter values are from each other. Specifically, we initially tested listeners with normal hearing. We gave these listeners a single 1–2 h training session to minimize any effects due to confusion about the task demands, and to account for any effects of rapid learning (e.g., Ortiz and Wright, 2009; Sabin *et al.*, 2012). We repeatedly used a single vocoded speech envelope to focus listeners on the amplitude envelope (by randomizing the carrier temporal fine structure), and to ensure that listeners had a rich template of the uncompressed signal. We also attempted to reduce the number of possible listening strategies by roving the presentation level (preventing a listening strategy based on overall level). We do note that this approach constrains us to a narrow set of testing conditions with somewhat low ecological validity. Therefore we also include two follow up tests of generality (Experiments IIa and IIb, described below).

After characterizing behavioral performance, we then attempted to identify acoustical features that correlated with

that performance. We focused on the acoustical features rather than the specific parameter values, with the hope that the resulting knowledge would be more generalizable. It is well known that the properties of the signal will interact with the compression (e.g., Stone and Moore, 1992). For instance, the temporal envelopes of speech sounds such as stop consonants (Turner *et al.*, 1992; van der Horst *et al.*, 1999) and affricates (Howell and Rosen, 1983) vary on time scales that would be distorted by compression. Indeed it has been observed that compression decreases consonant intelligibility, especially for high compression ratios and fast time constants (e.g., Jenstad and Souza, 2005). In contrast, compression will result in less distortion to the temporal envelopes of more steady-state features such as vowels. However, more spectral distortions can occur with multi-channel compression (e.g., Franck *et al.*, 1999, Bor *et al.*, 2008). Therefore, because the extent of compression-induced distortions is dependent upon the material sent to the compressor, the distortion produced by the same parameter values might be easily detectable for some input signals, but less detectable for others. With this in mind, here we attempted to quantify sensitivity to compression by analyzing the compressed signal itself, rather than focusing on the specific parameter value.

We also conducted two experiments (Experiments IIa and IIb) that were designed to begin to examine the extent to which the correlations that we identify in Experiment I extend beyond the initial testing paradigm. The first experiment was designed to examine performance on a new speech envelope that had more rapid fluctuations in the amplitude envelope (a 2× time compressed version of the original stimulus). The second experiment was designed to examine performance in a clinical population [listeners with sensorineural hearing impairment (HI)]. While these additional experiments still do not use ecologically valid stimuli, they do provide initial tests of generality.

The examination of HI listeners is particularly interesting not only because this population often wears hearing aids that use wide dynamic range compression, but also because performance on temporal perception tasks can be influenced by hearing loss. In particular, for individuals with sensorineural hearing loss, loudness grows more rapidly with increasing level, and therefore temporal fluctuations in amplitude could be magnified in their auditory system. This magnification could, in theory, improve the sensitivity of HI listeners to distortions to the temporal envelope, and there is some indication that such magnification can influence temporal perception (Glasberg *et al.*, 1987; Moore and Glasberg, 1988). However, when performance is assessed on a basic task such as ability to detect sinusoidal modulations to the amplitude envelope, there is usually little-to-no difference between HI listeners and those with normal hearing (Lamore *et al.*, 1984; Bacon and Viemeister, 1985; Bacon and Gleitman, 1992). It should be noted that such across-population comparisons are complicated by the fact that HI listeners require a greater sound pressure level to achieve the same sensation level, and that in most cases the audible bandwidth will be narrower in HI listeners. Increases in presentation level and bandwidth can

improve performance on temporal tasks (e.g., [Fitzgibbons, 1983](#); [Shailer and Moore, 1983](#); [Moore et al., 1992](#)). Thus, as long as the signal is loud enough and has a broad enough audible bandwidth, we would expect little difference between listeners with normal and impaired hearing. In the previous work examining discrimination of compression characteristics, HI listeners were reported to be more variable ([Gilbert et al., 2008](#)) or less sensitive ([Nabelek, 1984](#)) but note that they used different stimuli than listeners with normal hearing).

Finally, we also attempted to relate, on the individual level, the sensitivity to compression parameter values to performance on a basic psychoacoustic task. We chose to focus on modulation depth discrimination (e.g., [Wakefield and Viemeister, 1990](#); [Lee and Bacon, 1997](#)). In this task, listeners are asked to distinguish two amplitude-modulated noises that differ in terms of modulation depth. This is analogous to the compression task because compression reduces the difference between the peaks and valleys of the temporal envelope, effectively reducing the modulation depth ([Stone and Moore, 1992](#)). Accordingly, listeners perform more poorly on tests of temporal modulation tasks when the stimuli are compressed ([Brennan et al., 2013](#)). Here, we chose to examine modulation depth discrimination performance at 4 Hz, because speech has the most modulation energy at this frequency ([Houtgast et al., 1980](#); [Payton and Braida, 1999](#); [Holube et al., 2010](#)).

II. EXPERIMENT I

A. Overview

Listeners with normal hearing were tested in two sessions, each on a different day. During the first session, listeners were given practice on the dynamic range compression sensitivity and modulation depth discrimination tasks (both described below) for approximately one hour. Previous work has shown that a similar amount of practice (roughly 1–2 h) on a related task is sufficient to bring performance on a modulation depth-discrimination task to the asymptotic level ([Sabin et al., 2012](#)). During the second session, all listeners were tested on the same set of compression sensitivity conditions, with each condition comprised of trials that differed by a single compression parameter value. At the end of the second session, the listeners were tested on a sinusoidal modulation depth discrimination task. All testing used a single ear, which was arbitrarily chosen to be the left ear. Sounds were presented at a sampling rate of 44 100 Hz through the headphone amplifier of an external sound card (M-Audio Fasttrack) and an insert earphone (Etymotic ER-2).

B. Listeners

Twelve listeners (four females) who reported normal hearing completed the experiment. All listeners, were students at Northwestern University, and had a mean age of 22.1 yr (s.d. = 4.2 yr). All procedures were approved by the Institutional Review Board of Northwestern University, and all participants were compensated for their time.

C. Vocoding

The source material for the compression sensitivity task was a sentence processed with a single-channel vocoder (e.g., [Shannon et al., 1995](#)). The purpose of vocoding was to focus the listener on the temporal envelope (rather than the temporal fine structure), where the influence of compression is more likely to be apparent due to the relatively slow time range over which the compressor operates (see *Compression*, below). We do note however that in one other report ([Gilbert et al., 2008](#)) there was no difference in sensitivity to compression for vocoded vs normal speech. Here, the temporal envelope came from the arbitrarily selected single sentence “The picture came from a book,” which was presented at 65 dB sound pressure level (SPL). The sentence was 1.58 s in duration. Vocoding was accomplished by envelope extraction, and then multiplication of that envelope by a speech shaped noise ([Nilsson et al., 1994](#)). To extract the envelope we first half-wave rectified the time domain signal, and then passed the resulting signal through a fourth order Butterworth low pass filter with a cutoff frequency of 100 Hz. The envelope was multiplied by a random phase noise whose long-term average spectrum matched that of speech. The noise was created prior to each presentation, and thus had temporal fine structure that varied randomly from interval to interval (i.e., all three noises in a trial were independent). After vocoding, it was not possible for the listener to identify the words that made up the sentence.

D. Compression

Compression was applied offline using a custom-written algorithm. The compressor had five adjustable parameters: threshold (i.e., “knee point”), ratio, attack time constant, release time constant, and gain. All operations were performed on the temporal envelope, which was first extracted by computing the root-mean-square amplitude (in dB) with a 10 ms moving rectangular window ending with the current sample (the i th sample). This envelope was subsequently smoothed according to the attack and release time constants, as described below. The gain control signal was dependent on the smoothed temporal envelope input level (env , in dB SPL), the Gain ($Gain$, in dB), the Compression Threshold (Thr , in dB SPL) and the Ratio ($Ratio$, in dB/dB) according to the piecewise function described in Eq. (1).

$$Gain + \begin{cases} \frac{1 - Ratio}{Ratio}(env_i - Thr) & \text{if } env_i > Thr \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The input signal was delayed by 10 ms relative to the gain control signal. Such delays are commonly used to reduce the level of transient increases in sound level so that the control signal reduces the output right at the beginning of the transient (e.g., [Robinson and Huntington, 1973](#), [Verschuure et al., 1993](#)).

Before Eq. (1) was applied, the temporal envelope was smoothed in linear amplitude using the single pole low pass filter (IIR) described in Eq. (2).

$$env_i = (1 - \Theta)rms_i + \Theta env_{i-1}, \quad (2)$$

where rms is the temporal envelope extracted with the 10 ms moving rectangular window, and env is the smoothed temporal envelope. Note that the time constant Θ controls the influence of the previous smoothed envelope value on the current one. For instance, if $\Theta = 0$, then env is identical to rms . As Θ increases, the influence of a given envelope value remains for a longer time period, and thus more smoothing occurs.

Separate values of Θ were applied depending on whether the temporal envelope was increasing (in attack) or decreasing (in release) according to the Eq. (3)

$$\Theta \equiv \begin{cases} e^{-1/(F_s Att)} & \text{if } rms_i > env_i \\ e^{-1/(F_s Rel)} & \text{otherwise,} \end{cases} \quad (3)$$

where Att and Rel are the attack and release time constants and F_s is the sampling rate (in Hz) (as in Kates, 1993). We note that Att and Rel are not equivalent to the attack and release times as defined by ANSI (2003). To report values consistent with ANSI specification, we determined the attack and release times empirically. Specifically, the attack time was defined as the time it takes the output to drop to within 3 dB of the steady-state level after a 2000 Hz sinusoidal input changes from 55 dB SPL to 90 dB SPL. Release time was defined as time it takes the 2000 Hz sinusoidal output to stabilize to within 4 dB of the steady-state level after input changes from 90 to 55 dB SPL.

E. Compression sensitivity task and procedure

On each trial of the compression sensitivity task, the subject had to distinguish between a compressed and an uncompressed instance of the vocoded signal. Performance was evaluated using a cued three-interval, two-alternative force choice procedure with feedback. On each trial, three sounds were presented, and each presentation was marked by the appearance of a button on the computer screen. The first interval always contained an uncompressed signal. In the remaining two intervals, one of the sounds was compressed and the other was not. The presentation order of these two sounds was randomized. The listener was instructed to identify which of the three sounds had the least variation in loudness across its duration, and to ignore overall differences in presentation level between intervals. The listeners were shown graphic depictions of compressed and uncompressed amplitude envelopes. Before the first run of the compression sensitivity task, all listeners heard examples of uncompressed and compressed (with extreme values) signals. The listeners indicated their responses using a mouse to click on the appropriate box. Visual feedback indicating whether the listener's response was correct or incorrect immediately followed each response.

Performance was assessed using the method of constant stimuli where only one compression parameter varied within in a block of trials. We first defined a baseline set of compression parameter values (see Table I, top row). The baseline set was selected to be on the severe end of the values that might be prescribed in a clinical setting. In each trial within a block,

TABLE I. The tested compression parameter values for Experiment I.

Condition	Compression threshold (dB SPL)	Compression ratio	Attack time constant (Att)	Release time constant (Rel)
Baseline	50	2.7	0.001	0.04
Threshold	37.9	2.7	0.001	0.04
	57.4	2.7	0.001	0.04
	66.7	2.7	0.001	0.04
	70.8	2.7	0.001	0.04
Ratio	50	1.23	0.001	0.04
	50	1.5	0.001	0.04
	50	1.96	0.001	0.04
	50	2.91	0.001	0.04
Attack	50	2.7	0.004	0.04
	50	2.7	0.014	0.04
	50	2.7	0.027	0.04
	50	2.7	0.047	0.04
Release	50	2.7	0.001	0.002
	50	2.7	0.001	0.082
	50	2.7	0.001	0.246
	50	2.7	0.001	0.407

three of the four parameters applied to the compressed signal were taken from the baseline set. The value of the other parameter varied from trial-to-trial, and in many cases was more extreme than clinical norms. The set of tested parameter values was the same for every listener. The specific parameter values were selected using an initial acoustical analysis in which we identified the range of values over which the most substantial change to the temporal envelope occurred, defined according to the analysis described later in this paper. We selected four values that spanned each range (Table I). In a block of trials, each value was tested 10 times, for a total of 40 trials/block. Each parameter was tested in three blocks, thereby making 30 repetitions for each parameter value. Since there were four tested parameters, there were 12 blocks in total. Testing the entire set took approximately 90 min. The order of blocks was randomized across subjects.

Our goal was to measure sensitivity to the changes in temporal envelope shape introduced by compression. Accordingly, we took two steps to encourage listeners to ignore the absolute level of the stimuli. First, after compression, the Gain variable [see Eq. (1)] was set to scale each sound to have the same rms amplitude, which corresponded to a presentation level of 65 dB SPL (calibrated in a 2 cc coupler). We then recognized that the listeners could adopt an alternative strategy in which they select the interval that has the smallest peak temporal envelope value. To discourage this listening strategy, we randomized the presentation level of each sound ± 3 dB around the 65 dB SPL level (as is common in studies of profile analysis: Green, 1987). Together, these manipulations encouraged a listening strategy based on the relative shapes of the within-interval temporal envelopes rather than any absolute cue.

F. Modulation depth discrimination task and procedure

Listeners were also tested on their ability to discriminate the depth of sinusoidal temporal modulation. For this task,

the standard had 100% modulation depth ($m = 1$) and the signal had a shallower modulation depth ($m < 1$). The carrier was a 500 ms speech-shaped noise with 5 ms raised cosine on/off ramps. The modulator was sinusoidal in amplitude (ranging from 0 to 1), had a frequency of 4 Hz, and was always presented in cosine phase. The combination of modulation frequency and duration allowed for two full cycles of modulation on each trial. We used the same cued three-interval, two-alternative force choice task and the same ± 3 dB level randomization as the compression sensitivity task described above.

We estimated depth-discrimination thresholds using an adaptive two-down one-up procedure. The modulation depth of the signal increased following three correct answers and decreased following one incorrect answer. This procedure converges on the 70.7% correct point on the psychometric function (Levitt, 1971). A change in direction from increasing to decreasing, or vice versa, is called a reversal. Signal modulation depth was represented in terms of $20 \log_{10}(\Delta m)$ where Δm is the difference in modulation depth between the signal and standard (Lee and Bacon, 1997). The initial modulation depth was -10.7 dB—a value determined by pilot testing to yield tracks that regularly converged on a threshold. The step size was 1.1 dB until the third reversal, and 0.58 dB thereafter. In each block of 60 trials, the first three reversals were discarded, and the modulation depths at the largest remaining even number of reversals were averaged and taken as the modulation depth discrimination threshold. Each listener completed two blocks of modulation depth discrimination at the end of the testing session, taking approximately 10 min. In some cases a third block was completed if the performance across the first two blocks was not consistent.

G. Statistical analyses

Prior to any computation of a correlation coefficient, percent correct scores on the compression sensitivity task were transformed to rationalized arcsine units (RAU) (e.g., Sherbecoe and Studebaker, 2004). In all cases the Spearman rank correlation are reported.

H. Behavioral results

The individual psychometric functions for each tested compression parameter are plotted in left panels of Fig. 1. For eleven of the twelve tested subjects there was at least one set of compression parameters for which better than 90% correct performance was achieved. The performance of the remaining listener is plotted as a dashed line. The overall percent correct performance of this listener was more than two standard deviations from the mean of the other listeners, suggesting that this listener is an outlier. This listener also reported confusion about what to listen for. Therefore, the data from this listener were removed from analyses involving performance on the compression sensitivity task. We note that removal of this listener does not change any statistical conclusions based on significance, but removal does increase the strength of the correlations.

FIG. 1. Behavioral sensitivity to compression. (Left panels) The value of the manipulated compression parameter (x axis) vs the proportion correct discriminations between compressed and uncompressed signals. Results are plotted separately for each listener (each line), where the varied parameter was either (A) attack, (B) release, (C) threshold, or (D) ratio. In A and B, performance is shown in terms of ANSI time constants in milliseconds as well as Att or Rel from Eq. (3). The dotted line reflects the performance of an outlier, and the dashed line reflects chance performance. (Right panels) As in the corresponding left panel except that the average performance across all subjects (excluding the outlier) is plotted. Error bars reflect ± 1 standard error of the mean.

On the group level, sensitivity to compression varied monotonically with compressor parameter value, and was relatively consistent across subjects. The across-listener average performance is plotted in the right panels of Fig. 1. Sensitivity to compression increased as attack and release time constants decreased, as compression threshold decreased, and as compression ratio increased. For the stimuli and compression algorithm used here, each of these

directions was consistent with the direction of increasing amount of compression-induced distortion to the temporal envelope. We also note the relatively small across-listener variation. On average the standard deviation across all listeners at a given compression value was 9.9%.

There was also a relationship between performance on the compression sensitivity task and that on the modulation depth discrimination task. In Fig. 2 individual performance on the modulation depth discrimination task is plotted on the ordinate, and proportion correct across all compression sensitivity conditions is on the abscissa. There is a significant correlation in performance between these two tasks ($r = -0.68$, $p = 0.02$) such that listeners who performed better on the modulation depth discrimination task (lower thresholds) also did better on the compression sensitivity task (higher percent correct).

I. Comparison to acoustic features

We next set out to determine the extent to which the variations in behavioral performance across compression conditions were related to several acoustical features computed on the compressed signals. We first describe those acoustical features and then examine their relationship to behavioral performance. For each compression condition, we repeated the computation of each acoustical feature 50 times and took the average across repetitions in order to account for any variations due to the synthesis of the noise.

First, we considered various acoustical measures computed on the temporal modulation spectrum. We considered this feature because analysis of this modulation spectrum is incorporated into numerous successful models of auditory processing (Dau *et al.*, 1997b,a; Chi *et al.*, 1999; Chi *et al.*, 2005; Gallun and Souza, 2008; Jepsen *et al.*, 2008). Here, modulation frequency analysis was accomplished using the

FIG. 2. Relationship between modulation depth discrimination and sensitivity to compression. For each listener, the RAU-transformed percent correct performance on the compression sensitivity task (averaged across all parameter values) is on the x axis and the threshold for discriminating modulation depth is plotted on the y axis. The asterisk shows the performance of the outlier. Performance on these two tasks was significantly correlated such that better performance on the compression task (higher percent correct) was associated with better performance on the depth discrimination task (lower threshold).

method described by Gallun and Souza (2008). The time domain signal was half-wave rectified, and then passed through a fourth order low pass Butterworth filter with a 50 Hz cut-off. We then computed the fast Fourier transform of the envelope and averaged the magnitude in each of six modulation frequency bands with center frequencies of 1, 2, 4, 8, 16, and 32 Hz and a [3/4] octave bandwidth. Note that given the signal duration of 1.58 s it is possible to observe 1.5 cycles even at the lowest modulation frequency (1 Hz). All magnitude values were divided by the DC component of the modulation spectrum. We compared behavioral performance to the correlation between the modulation spectra of compressed and uncompressed signals [Spectral Correlation Index (SCI): Gallun and Souza, 2008; Souza and Gallun, 2010], and to the Euclidean distance between those two spectra. These computations were repeated using the modulation spectrum on both a linear amplitude scale and after conversion to a dB scale.

We also considered a global measure of the compression-induced reduction to the modulation depth. We first extracted the temporal envelope using the same procedure as described in the previous paragraph. We then estimated the global modulation depth by computing the standard deviation of the values comprising the temporal envelope. We compared behavioral performance to the difference in standard deviation between the compressed and uncompressed envelopes. We computed this value with the temporal envelope represented on a linear amplitude scale (Table II, left column) as well as on a dB scale (Table II, right column).

Finally, we also considered the envelope difference index (EDI)—a measure commonly used to quantify distortions to the temporal envelope (Fortune *et al.*, 1994). Briefly the EDI is the point-by-point average of the differences in the time-aligned envelopes of two signals. The EDI varies from zero (indicating that the two envelopes are identical) to one (indicating that the envelopes are opposite each other).

The Spearman correlation coefficients relating each of these acoustic features to the average behavioral performance are displayed in Table II. The strongest relationship to behavioral data was for the Euclidean distance between the modulation spectra for the compressed and uncompressed signals (Table II, row 2). A similarly strong correlation was found for the difference in modulation depth as estimated by

TABLE II. Relation to acoustic features. Each cell of the table is the Spearman rank correlation coefficient between an acoustic feature computed on the stimuli used in the compression sensitivity task, and the across-subject average performance on that task.

Measure	Correlation to average behavioral data	
	Linear scale	dB scale
Modulation spectrum correlation (SCI)	−0.41	−0.47
Modulation spectrum Euclidean distance	0.91 ^a	0.90 ^a
Envelope standard deviation	0.93 ^a	0.90 ^a
Envelope difference index	0.55 ^b	n/a

^a $p < 0.0001$.

^b $p < 0.05$.

the difference in envelope standard deviation between the compressed and uncompressed signals (Table II, row 3). For both measures, there was little difference in the strength of this correlation whether the feature was represented on a linear (Table II, left column) or (Table II, right column) scale. For the SCI (Table II, row 1) there was no correlation to the behavioral data. Finally, for the EDI (Table II, row 4) there was a slight, but far less significant, correlation.

We further examined the correlation to modulation spectrum Euclidean distance on a modulation-frequency-specific basis. We repeated the same Euclidean distance based analysis described above (linear scale), but computed each correlation separately for each of the six modulation frequency bands [Fig. 3(A)]. The strongest correlation between Euclidean distance and behavioral performance occurred at the 4 Hz band

($r = 0.93$, $p < 0.0001$) and decreased as the examined modulation frequency got further away from 4 Hz. Note that the correlation coefficient for the 4 Hz band ($r = 0.93$) is even higher than that computed from all bands combined ($r = 0.91$) suggesting that the inclusion of the other bands did not improve the metric, or may have even made it worse.

It appears that the dominance of the 4 Hz modulation band in the correlation analysis might simply be due to the large magnitude of modulation at that frequency. The modulation spectrum of the uncompressed signal used in the compression sensitivity task is plotted in Fig. 3(B). The modulation spectrum has a peak at 4 Hz. This is consistent with the regular observation that speech has the highest magnitude at this modulation frequency (Houtgast *et al.*, 1980; Payton and Braida, 1999; Holube *et al.*, 2010), due to the syllable rate.

FIG. 3. Modulation frequency dependence. (A) Summary of the modulation spectrum distance vs compression sensitivity. Spearman rank correlation coefficients, computed separately for each modulation frequency. The correlation was strongest at 4 Hz and decreased for lower and higher modulation frequencies. (B) The modulation spectrum of the standard used in the compression sensitivity tasks. The behavior vs modulation distance correlation was highest at the peak of the modulation spectrum.

III. EXPERIMENT II

The generality of the correlation to acoustic features observed in Experiment I was examined in two follow-up experiments in which performance was assessed either in the same population, but with a different stimulus (Experiment IIa), or with the same stimulus, but in a population of listeners with sensorineural hearing loss (Experiment IIb).

A. Method, Experiment IIa

The method used in Experiment IIa was nearly identical to that in Experiment I. Only differences from Experiment I will be described here. The test subjects included a subset of six of the normal hearing listeners who had first participated in Experiment I.

The primary difference from Experiment I was the use of different test stimulus. In Experiment IIa, the test stimulus was the one used in Experiment I (a man speaking “the picture came from a book” passed through a signal channel noise vocoder) but time-compressed by a factor of 2. Time compression was accomplished by selecting every other sample from the original (non-vocoded) recording. We chose to test this condition for two reasons: first, to begin to examine whether the correlation to acoustic features changes with a different stimulus, and second to effectively shift the modulation spectrum up one octave to identify if the pattern of correlations between behavioral performance and frequency-specific modulation distortion would follow this shift.

B. Results, Experiment IIa

Analyses of Experiment IIa focused on comparisons of behavioral performance between the normal and time compressed stimuli. As before, our goal was to relate performance to the acoustical features of the influence of compression, rather than the compression parameters themselves. We focus on the Euclidean distance between the modulation spectra of compressed and uncompressed signals because we observed the highest correlation with that feature.

The normal and time-compressed stimuli led to different amounts of compression-induced distortion for the same set of parameter values, and therefore direct comparisons of percent correct for the same distortion level are not possible.

Instead, we inferred percent correct values across a broad range of distortions for each listener in Experiment I by fitting a psychometric function to the modulation distance vs percent correct function. Fitting was accomplished using a psychometric function fitting toolbox (Prins and Kingdom, 2009) with the two parameter logistic function.

$$y = 0.5 + 0.5 \frac{1}{1 + e^{-\beta(x-\alpha)}}, \quad (4)$$

where x is the modulation spectrum distance, y is the estimated percent correct, and α and β are the two fitted parameters that control the function horizontal position and slope, respectively. We computed the 95% confidence interval of the mean psychometric function computed across all listeners in Experiment I. This interval is displayed in Fig. 4(A) as the filled gray area.

FIG. 4. Performance of normal hearing listeners using a time-compressed signal. (A) The gray filled area represents the 95% confidence interval of the across listener mean psychometric function (modulation spectrum distance vs proportion correct) using the normal (not time compressed) stimulus from Experiment I. The circles represent the mean performance of a subset of six listeners with normal hearing using a 2 $\times$ -time-compressed version of the stimulus from Experiment I. Error bars are ± 1 standard error of the mean. Both data sets show a similar correlation. (B) As in Fig. 3(A) except plotted for the time-compressed data (solid line with circles). For comparison the correlation function from Experiment I is plotted as well (dotted line with squares). The pattern of correlation vs modulation shifted upward for the time compressed signal.

Sensitivity to dynamic range compression for the time-compressed stimulus was similar to that for the normal stimulus. The circles in Fig. 4(A) represent the across-listener average performance for the time compressed stimulus. As in the original experiment, there was a strong and significant correlation between modulation spectrum distance and behavioral performance ($r = 0.90$, $p < 0.0001$). Further, note that for nearly all distortion levels there is some overlap between the range of performance on the time compressed condition and the normal condition (the overlap between the circles and the gray fill). We compared the fitted psychometric function parameters between conditions (time compressed vs normal) using two-group t -tests. Suggesting that there were no systematic differences in performance between the conditions, there were no differences in α ($t_{15} = -1.26$, $p = 0.23$) or β ($t_{15} = 1.17$, $p = 0.26$). We do note that it does appear that in general performance on the time compressed condition falls on the lower range of performance of the normal condition, but that is not a statistically significant difference. Thus the observed correlation to modulation spectrum distance is not restricted to the test stimulus used in Experiment I.

We also examined whether the correlation to modulation-frequency-specific distortions differed between the normal and time-compressed conditions. We reasoned that if the strength of correlation was related to the modulation energy, then the function relating modulation frequency to correlation coefficient would shift up an octave for the time-compressed signal. Alternatively, if the listeners had a bias to attend to modulations at 4 Hz, the shape of the modulation frequency vs correlation function would not change with the time-compressed signal. This function for the time-compressed signal is displayed in Fig. 4(B) (solid line with circles). The 1 Hz point is omitted from this graph because the duration of the stimulus (0.79 s) was not sufficient to capture a full cycle at this frequency. For reference, the same function is re-plotted for the original (non-time-compressed) analysis (dashed line with squares). The peak of the time-compressed function is now at 8 Hz, indicating an upward shift of one octave. Thus, it appears the correlation at a given modulation frequency was related to the modulation energy at that frequency.

C. Method, Experiment IIb

As for Experiment IIa, here we will only describe differences between Experiments I and IIb. The primary difference was in population of tested listeners. Here we tested seven listeners with sensorineural hearing impairment (see Table III for listener demographics and audiograms). Prior to testing, the hearing of each listener was assessed with air and bone conduction audiograms, as well as tympanometry. In all cases the air-bone gap was ≤ 10 dB and the tympanometry was within lab norms, suggesting a sensorineural hearing loss. These listeners were selected to be consistent with a typical clinical population.

Pilot testing indicated that there was some confusion in this population regarding the goal of the task when using the same parameters as in Experiment I. To address this issue,

TABLE III. Listener demographics and audiograms from Experiment IIb. Gender, age (in years), and left (tested) ear threshold (in dB HL) of each of the listeners with sensorineural hearing loss.

Subject	Gender	Age	Frequency (kHz)							
			0.25	0.5	1	2	3	4	6	8
24	M	76	15	25	10	25	35	60	70	70
33	F	67	30	35	25	35	45	60	60	55
47	M	66	20	20	20	35	35	55	70	75
48	F	81	20	40	60	65	60	60	60	65
51	F	74	20	25	30	40	45	45	70	85
53	F	81	45	45	40	45	50	50	50	65
68	M	66	20	5	10	25	35	45	50	45

the HI listeners were tested on a range of parameter values that was designed to create distortions that were more severe than those from Experiment I. This served to make the object of the task more clear. These more severe testing parameter values are displayed in Table IV. Further, each trial took nearly twice as long in these listeners in comparison to the normal hearing listeners. To make the experiment duration comparable between groups, we tested listeners with hearing impairment on eight different compression conditions (in comparison to the 16 conditions in the normal hearing listeners).

Stimulus synthesis was the same as in Experiment I, except that a listener-specific frequency-gain curve was applied to each stimulus in the compression sensitivity and modulation-depth discrimination tasks. The frequency gain curve was derived using the NAL-RP formula (Byrne and Dillon, 1986). The resulting frequency vs gain curve was implemented using a single linear phase finite impulse response (FIR) filter. Finally, to accommodate the increased sound level, sounds were output from the computer through a different digital to analog converter and headphone amplifier (RX6 and HB7, Tucker-Davis Technologies System 3) before they were sent to the insert earphone.

D. Results, Experiment IIb

As in Experiment IIa, the primary focus of analysis was a comparison to the results from Experiment I. Again, a

TABLE IV. The parameters used for the simulations presented in the figures, except where otherwise noted.

Condition	Compression threshold (dB SPL)	Compression ratio	Attack time constant (Att)	Release time constant (Rel)
Baseline	50	5	0.001	0.001
Threshold	62.2	5	0.001	0.001
	49.9	5	0.001	0.001
Ratio	50	2.2	0.001	0.001
	50	2.5	0.001	0.001
Attack	50	5	0.003	0.001
	50	5	0.093	0.001
Release	50	5	0.001	0.028
	50	5	0.001	0.219

FIG. 5. Performance of listeners with hearing impairment. (A) As in Fig. 4(A), except plotted for a group of listeners with sensorineural hearing loss (triangles). Both data sets show a similar correlation, though the listeners with hearing impairment show a trend toward a steeper, but otherwise similar, psychometric function. (B) Individual performance on the modulation depth discrimination task for the listeners with normal (left, squares) and impaired (right, triangles) hearing. Individual points horizontally staggered for ease of viewing. The box is next to the individual points is comprised of lines at the upper quartile, median, and lower quartile values, and the whiskers extend from each end of the box to the maximum and minimum values for that group. Performance on this task was significantly worse for the listeners with hearing impairment.

direct comparison on the identical conditions was not possible because the listeners with HI were not tested using the same compression settings as those with normal hearing (NH). As above, performance of the NH group at the compression-induced distortion levels tested in the current experiment was inferred by fitting psychometric functions.

Performance on the compression sensitivity task was largely similar between the NH and HI groups. The average performance of the HI group is plotted in Fig. 5(A) (triangles) with the 95% confidence interval of the mean performance of the NH group in Experiment I. As in the previous experiments, behavioral performance is significantly correlated to the Euclidean distance between the modulation spectra of the compressed and uncompressed signals ($r=0.79$, $p=0.02$). For nearly all points, the performance of the HI group fell well within the range of performance of the NH

group. As above, performance between groups was evaluated by comparing the fitted parameters of the psychometric functions. While there was no difference between groups in α (horizontal position: $t_{16} = 1.13$, $p = 0.277$), there was a trend toward a significant difference in β (slope; $t_{16} = 2.01$, $p = 0.06$). Thus it appears that the NH and HI groups had similar psychometric functions, but the HI listeners might have had a steeper psychometric function in which performance worsened at a quicker rate as the compression-induced distortion got smaller.

Finally, the HI group performed more poorly than the NH group on the modulation depth discrimination task. The performance of the HI listeners (right, triangles) on the depth discrimination task is plotted next to that of the NH listeners (left, squares) in Fig. 5(B). On the group level, HI listeners performed significantly worse than NH listeners ($t_{16} = -3.75$, $p < 0.01$) by an average difference of 6.7 dB. Providing an additional contrast to the NH listeners, for the HI listeners there was no correlation between the threshold for modulation depth discrimination and the overall performance on the compression sensitivity task ($r = 0.04$, $p = 0.94$).

IV. DISCUSSION

Listeners with NH were consistently sensitive to compression parameters, and that sensitivity varied monotonically with compression parameter value in a predictable way. On average, performance was highly correlated to several acoustical indices of the compression-induced change in overall modulation depth, especially the Euclidean distance between the modulation spectra of the compressed and uncompressed signals (Experiment I). This correlation remained largely unchanged when normal hearing listeners were tested with a time-compressed signal (Experiment IIa) or when a group of listeners with HI were tested with the normal stimulus, but with an individualized frequency vs gain curve applied (Experiment IIb). On the individual level, sensitivity to compression was related to performance on a 4 Hz modulation depth discrimination task in listeners with normal hearing, but not in listeners with impaired hearing. Listeners with impaired hearing were significantly worse on the depth-discrimination test.

One difference in behavioral results between the current and previously published work is that we observed less individual variability. In the two investigations described in the introduction (Nabelek, 1984; Gilbert *et al.*, 2008), a central observation was that there was a large amount of individual-to-individual variation in sensitivity to compression parameters ranging from chance to perfect performance on most of the measured conditions. In contrast, here nearly all NH listeners were highly sensitive to some compression conditions. Eleven of the twelve NH listeners reached at least 90% correct on at least one condition. Here, because our goal was to characterize performance under highly favorable listening conditions we took several steps that could have reduced individual variability. First, we gave all listeners roughly one hour of practice on the compression sensitivity task on a day preceding testing. A single one to two hour session of practice could have brought listeners to optimal performance on

a depth-discrimination task (Sabin *et al.*, 2012). Next, because the listeners in Experiment I were all younger university students it is possible that they all had high cognitive abilities. Previous work has shown that listeners with high cognitive abilities show more benefit to changes in compression parameters such as the inclusion of fast time constants (Gatehouse *et al.*, 2003). Additionally, the same vocoded signal was used on every trial. This factor likely allowed the listeners to develop a rich internal template of the uncompressed signal, making the deviations from that template easy to detect. The vocoding might have encouraged the listeners to focus on the amplitude envelope. Finally, unlike the previous work, we added an overall level randomization to the task. As mentioned in the method of Experiment I, even after the overall presentation level was normalized, the subject could have based their response on the interval that had the smallest peak of the amplitude envelope. To reduce the likelihood of this listening strategy we randomized the overall level over a 6 dB range. It is possible that in previous work, part of the individual variability was attributable to some listeners adopting a strategy based on the relative shapes of the temporal envelopes and others adopting one based on the absolute level of the temporal envelope peak.

On the population level, sensitivity to a particular compression setting is highly correlated to differences between acoustical features that summarize overall change in modulation depth caused by the compressor. Behavioral performance was highly correlated to the Euclidean distance between compressed and uncompressed modulation spectra, as well as to the compressed vs uncompressed difference in temporal envelope standard deviation. This sensitivity was not well predicted by the correlation between two modulation spectra (SCI: Gallun and Souza, 2008; Souza and Gallun, 2010). This correlation was likely not predictive because compression did not change the relative shape of the modulation spectrum, but reduced the difference between the deepest and shallowest modulations. Thus, because compression did not change relative shape of the modulation spectrum, the correlation between compressed and uncompressed modulation spectra remained high even for large amounts of compression. The SCI might be more effective at accounting for confusions between signals, rather than different distortions to the same signal. Similarly, the time-aligned EDI was not a good predictor of compression sensitivity. This might be due to the fact that the EDI is a point-by-point comparison between the compressed and uncompressed envelopes, rather than a comparison of summary statistics of those envelopes.

While Experiment I tested a narrow range of conditions (to characterize performance under highly favorable listening conditions), the results from Experiment IIa give some indication that the correlation to acoustic features are not restricted to the test stimulus. The correlation between modulation spectrum distance and behavioral performance remained largely unchanged when a subset of the NH group performed the same task using a time-compressed stimulus [Fig. 4(A)]. This result implies that the correlation between behavior and modulation spectrum distance remains after a dramatic change in the modulation spectrum to signal that

is input to the compressor. Further, the results from the time-compressed condition indicate that the pattern of that correlation across modulation frequency [Fig. 4(B)] is tied to the modulation spectrum of the uncompressed stimulus, rather than a consistent bias toward a particular modulation frequency. Overall, we consider the time-compressed condition to be an initial suggestion of generality, but further tests of generality should examine non-vocoded signals and a variety of natural sounds. Such test should also examine multiband compression, where co-modulation between bands could provide an additional cue for listeners (e.g., Stone and Moore, 2004). These questions continue to be a focus of interest in our laboratory.

The results of Experiment IIb provide some indication that these correlations between modulation spectrum distance and behavioral performance extend to listeners with impaired hearing. In general, performance of the HI group on the compression sensitivity task was correlated to the distance in modulation spectra between the compressed and uncompressed signals [Fig. 5(A)], and was of similar performance level to the NH group. However, there was some indication that the HI listeners had steeper psychometric function slope than the NH listeners. If so, then this difference might contribute to the large individual variability that has been reported for HI listeners on tests of sensitivity to dynamic range compression (Nabelek, 1984; Gilbert *et al.*, 2008). A small change to the influence of the compressor (either by changing the input level or compression parameter) can lead to a large change in performance if the change is made where the slope of the psychometric function is steepest. We do note that the HI group was also older than the NH group. Thus any differences between groups could potentially be attributed to age. However, we do note that, on the individual level, performance on the compression sensitivity task was not correlated to age or pure tone average threshold (all $p > 0.05$).

We also expand upon previous findings to show that, for NH listeners, sensitivity to compression is related to an individual's ability to discriminate the depth of sinusoidal temporal modulation (Fig. 2). This relationship is understandable given the similarity between the two tasks. Namely, the overall effect of the compressor was to reduce the depth of the temporal modulations in the original signal. Thus, both tasks force the subjects to distinguish sounds based on their modulation depths, and both tasks used modulated noise. Interestingly, this correlation was not observed in the HI group, who on average, were significantly worse at depth discrimination than the NH group [Fig. 5(B)].

One possible source of difference in overall performance between the NH and HI groups could be related to the across-population differences in psychometric function slope described in the previous paragraph. The modulation depth discrimination task used a two-down one-up procedure which should converge on the 70.7% correct point on the psychometric function (Levitt, 1971). If the psychometric function is truly steeper, but otherwise similar, for the HI group, then the differences in performance between groups should only emerge when performance is tested at a difficult level, such as the 70.7% correct point. This explanation

could account for why depth discrimination performance was worse in the HI group (where performance was tested at this difficult level) but was largely similar between groups on the compression sensitivity task (where performance was tested at easier levels). This account does not explain why performance on the two tasks was correlated in NH but not in HI listeners. Unlike NH listeners, in HI listeners it appears that performance on the two tasks is not limited by the same or similar factors.

It is important to recognize that the correlations to acoustic features described here can account for cases in which compression-induced changes to overall loudness are ignored. In all experiments in the current work, compressed stimuli were scaled to have the same *rms* amplitude that they had before compression. In hearing aids, compression is used primarily to adjust loudness. Specifically, the ultimately effect of compression is to reduce the gain applied to high input level sounds. A more complete model of sensitivity to compression would also need to take into account overall loudness perception.

Finally, this work can ultimately be used to constrain the rather large parameter space of a dynamic range compressor. If we assume that each of the four parameters can take on 20 values (these values are chosen arbitrarily), then there are a total of 160 000 (20^4) possible parameter value combinations. As described previously, we characterized the sensitivity to compression-induced distortions under highly favorable listening conditions. Thus, this work could potentially be used to identify the subset of parameter value combinations that are barely distinguishable by most listeners under highly favorable listening conditions (e.g., combinations for with the modulation difference would predict a lower percent correct), thus providing a principled way to reduce the overall parameter space. However, this possibility is contingent on how well the relationships described here generalize to more ecologically valid conditions involving natural sounds and multiband compression.

ACKNOWLEDGMENTS

Efoe Nyatepe-Coo and Alex Evanoff helped with data collected. Morten Jepsen, Eric Hoover and Richard Wright provided helpful conversations in the development of this work. This work was funded by NIH grant DC60014 and by the Department of Veterans Affairs, Veterans Health Administration, Rehabilitation Research and Development Service (Center of Excellence Award C4844C).

- ANSI (2003). S3.22-2003, *Specification of Hearing Aid Characteristics* (Acoustical Society of America, New York).
- Bacon, S. P., and Gleitman, R. M. (1992). "Modulation detection in subjects with relatively flat hearing losses," *J. Speech Hear. Res.* **35**, 642–653.
- Bacon, S. P., and Viemeister, N. F. (1985). "Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners," *Audiology* **24**, 117–134.
- Bor, S., Souza, P., and Wright, R. (2008). "Multichannel compression: effects of reduced spectral contrast on vowel identification," *J. Speech Lang. Hear. Res.* **51**, 1315–1327.
- Brennan, M. A., Gallun, F. J., Souza, P. E., and Stecker, C. (2013). "Temporal resolution with a prescriptive fitting formula," *Am. J. Audiol.* (in press).

- Byrne, D., and Dillon, H. (1986). "The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid," *Ear Hear.* **7**, 257–265.
- Chi, T., Gao, Y., Guyton, M. C., Ru, P., and Shamma, S. (1999). "Spectrotemporal modulation transfer functions and speech intelligibility," *J. Acoust. Soc. Am.* **106**, 2719–2732.
- Chi, T., Ru, P., Shamma, S. A. (2005). "Multiresolution spectrotemporal analysis of complex sounds," *J. Acoust. Soc. Am.* **118**, 887–906.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997a). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," *J. Acoust. Soc. Am.* **102**, 2892–2905.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997b). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," *J. Acoust. Soc. Am.* **102**, 2906–2919.
- Fitzgibbons, P. J. (1983). "Temporal gap detection in noise as a function of frequency, bandwidth, and level," *J. Acoust. Soc. Am.* **74**, 67–72.
- Fortune, T. W., Woodruff, B. D., and Preves, D. A. (1994). "A new technique for quantifying temporal envelope contrasts," *Ear Hear.* **15**, 93–99.
- Franck, B. A., van Kreveld-Bos, C. S., Dreschler, W. A., and Verschuure, H. (1999). "Evaluation of spectral enhancement in hearing aids, combined with phonemic compression," *J. Acoust. Soc. Am.* **106**, 1452–1464.
- Gallun, F., and Souza, P. (2008). "Exploring the role of the modulation spectrum in phoneme recognition," *Ear Hear.* **29**, 800–813.
- Gatehouse, S., Naylor, G., and Elberling, C. (1999). "Aspects of auditory ecology and psychoacoustic function as determinants of benefit from and candidature for non-linear processing in hearing aids," in *Auditory Models and Non-Linear Hearing Instruments*, edited by A. N. Rasmussen, P. A. Osterhammel, T. Anderson, and T. Poulsen, pp. 221–233: (Holmens Trykkeri, Copenhagen, Denmark), pp. 221–233.
- Gatehouse, S., Naylor, G., and Elberling, C. (2003). "Benefits from hearing aids in relation to the interaction between the user and the environment," *Int. J. Audiol.* **42** Suppl. 1, S77–85.
- Gatehouse, S., Naylor, G., and Elberling, C. (2006). "Linear and nonlinear hearing aid fittings—I. Patterns of benefit," *Int. J. Audiol.* **45**, 130–152.
- Gilbert, G., Akeroyd, M. A., and Gatehouse, S. (2008). "Discrimination of release time constants in hearing-aid compressors," *Int. J. Audiol.* **47**, 189–198.
- Glasberg, B. R., Moore, B. C. J., and Bacon, S. P. (1987). "Gap detection and masking in hearing-impaired and normal-hearing subjects," *J. Acoust. Soc. Am.* **81**, 1546–1556.
- Green, D. M. (1987). *Profile Analysis: Auditory Intensity Discrimination* (Oxford University Press, Oxford, UK), pp. 1–144.
- Hickson, L. (1994). "Compression amplification in hearing aids," *Am. J. Audiol.* **3**, 51–65.
- Holube, I., Fredelake, S., Vlaming, M., and Kollmeier, B. (2010). "Development and analysis of an International Speech Test Signal (ISTS)," *Int. J. Audiol.* **49**, 891–903.
- Houtgast, T., Steeneken, H., and Plomp, R. (1980). "Predicting speech intelligibility in rooms from the modulation transfer function. I. general room acoustics," *Acustica* **46**, 60–72.
- Howell, P., and Rosen, S. (1983). "Production and perception of rise time in the voiceless affricate/fricative distinction," *J. Acoust. Soc. Am.* **73**, 976–984.
- Jenstad, L. M., and Souza, P. E. (2005). "Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility," *J. Speech Lang. Hear. Res.* **48**, 651–667.
- Jepsen, M. L., Ewert, S. D., and Dau, T. (2008). "A computational model of human auditory signal processing and perception," *J. Acoust. Soc. Am.* **124**, 422–438.
- Kates, J. M. (1993). "Optimal estimation of hearing-aid compression parameters," *J. Acoust. Soc. Am.* **94**, 1–12.
- Kates, J. M., and Arehart, K. (2010). "The Hearing-Aid Speech Quality Index (HASQI)," *J. Audio Eng. Soc.* **58**, 363–381.
- Lamore, P. J., Verweij, C., and Brocaar, M. P. (1984). "Reliability of auditory function tests in severely hearing-impaired and deaf subjects," *Audiology* **23**, 453–466.
- Lee, J., and Bacon, S. P. (1997). "Amplitude modulation depth discrimination of a sinusoidal carrier: effect of stimulus duration," *J. Acoust. Soc. Am.* **101**, 3688–3693.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Moore, B. C. J. (2008). "The choice of compression speed in hearing aids: theoretical and practical considerations and the role of individual differences," *Trends Amplif.* **12**, 103–112.
- Moore, B. C. J., and Glasberg, B. R. (1988). "Gap detection with sinusoids and noise in normal, impaired, and electrically stimulated ears," *J. Acoust. Soc. Am.* **83**, 1093–1101.
- Moore, B. C. J., Shailer, M. J., and Schooneveldt, G. P. (1992). "Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss," *Br. J. Audiol.* **26**, 229–237.
- Nabelek, I. V. (1984). "Discriminability of the quality of amplitude-compressed speech," *J. Speech Hear. Res.* **27**, 571–577.
- Nilsson, M., Soli, S. D., Sullivan, J. A. (1994). "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.* **95**, 1085–1099.
- Ortiz, J. A., and Wright, B. A. (2009). "Contributions of procedure and stimulus learning to early, rapid perceptual improvements," *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 188–194.
- Payton, K. L., and Braida, L. D. (1999). "A method to determine the speech transmission index from speech waveforms," *J. Acoust. Soc. Am.* **106**, 3637–3648.
- Prins, N., and Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data. <http://www.palamedestoolbox.org> (date last viewed 4/30/13).
- Robinson, C. E., and Huntington, D. A. (1973). "The intelligibility of speech processed by delayed long-term averaged compression amplification," *J. Acoust. Soc. Am.* **54**, 314.
- Sabin, A. T., Eddins, D. A., and Wright, B. A. (2012). "Perceptual learning evidence for tuning to spectrotemporal modulation in the human auditory system," *J. Neurosci.* **32**, 6542–6549.
- Shailer, M. J., and Moore, B. C. J. (1983). "Gap detection as a function of frequency, bandwidth, and level," *J. Acoust. Soc. Am.* **74**, 467–473.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Sherbecoe, R. L., and Studebaker, G. A. (2004). "Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units," *Int. J. Audiol.* **43**, 442–448.
- Souza, P. E. (2002). "Effects of compression on speech acoustics, intelligibility, and sound quality," *Trends Amplif.* **6**, 131–165.
- Souza, P., and Gallun, F. (2010). "Amplification and consonant modulation spectra," *Ear Hear.* **31**, 268–276.
- Stone, M. A., and Moore, B. C. J. (1992). "Syllabic compression: effective compression ratios for signals modulated at different rates," *Br. J. Audiol.* **26**, 351–361.
- Stone, M. A., and Moore, B. C. J. (2004). "Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task," *J. Acoust. Soc. Am.* **116**, 2311–2323.
- Stone, M. A., Moore, B. C. J. (2007). "Quantifying the effects of fast-acting compression on the envelope of speech," *J. Acoust. Soc. Am.* **121**, 1654–1664.
- Stone, M. A., Moore, B. C. J., Alcantara, J. I., and Glasberg, B. R. (1999). "Comparison of different forms of compression using wearable digital hearing aids," *J. Acoust. Soc. Am.* **106**, 3603–3619.
- Turner, C., Horwitz, A., and Souza, P. (1992). "Identification and discrimination of stop consonants: Formants versus spectral peaks," in *Auditory Physiology and Perception*, edited by Y. Cazals *et al.* (Pergamon, Oxford, UK), pp. 463–470.
- van der Horst, R., Leeuw, A. R., and Dreschler, W. A. (1999). "Importance of temporal-envelope cues in consonant recognition," *J. Acoust. Soc. Am.* **105**, 1801–1809.
- Verschuure, J., Dreschler, W. A., de Haan, E. H., van Cappellen, M., Hammerslag, R., Mare, M. J., Maas, A. J., and Hijmans, A. C. (1993). "Syllabic compression and speech intelligibility in hearing impaired listeners," *Scand. Audiol. Suppl.* **38**, 92–100.
- Wakefield, G. H., and Viemeister, N. F. (1990). "Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise," *J. Acoust. Soc. Am.* **88**, 1367–1373.