Acoustic Analysis of Speech through a Hearing Aid: Perceptual Effects of Changes with Two-Channel Compression

Louise Hickson*
Nick Thyer††

Abstract
Compression amplification significantly alters the acoustic speech signal in comparison to linear amplification. The central hypothesis of the present study was that the compression settings of a two-channel aid that best preserved the acoustic properties of speech compared to linear amplification would yield the best perceptual results, and that the compression settings that most altered the acoustic properties of speech compared to linear would yield significantly poorer speech perception. On the basis of initial acoustic analysis of the test stimuli recorded through a hearing aid, two different compression amplification settings were chosen for the perceptual study. Participants were 74 adults with mild to moderate sensorineural hearing impairment. Overall, the speech perception results supported the hypothesis. A further aim of the study was to determine if variation in participants' speech perception with compression amplification (compared to linear amplification) could be explained by the individual characteristics of age, degree of loss, dynamic range, temporal resolution, and frequency selectivity; however, no significant relationships were found.

Key Words: Acoustic analysis, compression amplification, linear amplification, speech perception

Abbreviations: CV = consonant-vowel; CVR = consonant-vowel ratio; EDI = envelope difference index; LGOB = loudness growth in octave bands; MCL = most comfortable level; rms = root-mean-square; SPL = sound pressure level; TB = threshold masked by a broad band noise; TC = threshold in continuous noise; TM = threshold in continuous modulated noise; TN = threshold masked by a notched noise; TQ = threshold in quiet; TR = temporal resolution

Sumario
La amplificación por compresión altera significativamente la señal acústica del lenguaje en comparación con la amplificación lineal. La hipótesis central del presente estudio consistió en asumir que en un audífono de dos canales, los ajustes de compresión que preservan mejor las propiedades acústicas del lenguaje, proporcionarán los mejores resultados de percepción, al compararlos con los de una amplificación lineal. Asimismo, los ajustes de compresión que más alteran las propiedades acústicas del lenguaje, comparados con la amplificación lineal, generarán una percepción del lenguaje significativamente más pobre. Basados en el análisis acústico inicial de los estímulos de prueba registrados por medio de un auxiliar auditivo, dos grupos diferentes de ajustes para la amplificación de compresión fueron escogidos para el estudio de percepción. Los participantes fueron 74 adultos con trastornos auditivos sensorineurales de grado leve a moderado. En general, los resultados obtenidos de percepción del lenguaje apoyaron la hipótesis. Un propósito adicional del estudio fue determinar si las variaciones en la percepción del lenguaje de los...
E xperimental studies on speech perception with multichannel compression have been characterized by their variability. Some studies have shown positive benefits of compression compared to conventional linear amplification (Benson et al., 1992; Moore et al., 1992; Kam and Wong, 1999; Souza and Turner, 1999); others have reported no difference (Plomp, 1994; Crain and Yund, 1995; van Harten de Bruijn et al., 1997), and there are those that have reported detrimental effects of compression on speech perception for some individuals (Crain and Yund, 1995; Hickson et al., 1995; Hornsby and Ricketts, 2001; Plomp, 1994). It has been suggested that closer examination of the acoustic properties of the compressed speech signal may help to explain some of the variation in findings (Hickson, 1994; Plomp, 1994; Hickson et al., 1999a).

With single-channel compression, significant acoustic changes are introduced into the speech signal that are different from those seen in linear amplification systems. The most commonly reported change is an increase in the consonant-vowel ratio (CVR), where the intensity of consonants is increased and the intensity of vowels is decreased (Hickson and Byrne, 1995; Sammeth et al., 1996). CVR changes may also occur with linear amplification that has a high-frequency emphasis, and the extent of any CVR change depends on a combination of the compression characteristics and the frequency-gain response. There is evidence that CVR may be a cue for the perception of some consonant sounds for people with hearing impairment, and therefore increasing the CVR is not always beneficial (Hickson et al., 1995; Hickson and Byrne, 1997; Hedrick and Rice, 2000). It is therefore likely that decisions regarding selection of compression characteristics, when fitting a listener’s dynamic range, are also likely to affect their speech perception. Changes to the speech signal with multichannel compression amplification are likely to be more complicated than the effects of single-channel compression, and a number of researchers have suggested that spectral distortion caused by multichannel compression adversely affects speech perception (Moore, 1991; Plomp, 1994; Crain and Yund, 1995). Thus, the primary aim of the present study was to investigate acoustic changes to the speech signal that occurred with two-channel compression and to examine the relationship between such changes and speech perception by people with sensorineural hearing impairment.

In a previous article, we reported the CVR effects of varying compression ratio and crossover frequency with two-channel syllabic compression amplification (Hickson et al., 1999a). Eight voiceless consonants, each paired with three vowels, were processed by 12 types of two-channel compression amplification and one linear amplification condition. Acoustic analysis of the 24 consonant-vowel syllables indicated that compression in the low-frequency channel paired with linear amplification in the high-frequency channel was associated with little change in CVR compared with linear amplification. Higher CVRs were obtained with compression in the high-frequency channel and when the crossover frequency between channels was increased. The present
article reports on additional acoustic analysis of these same speech stimuli and goes a step further by examining the perceptual effects of the acoustic changes caused by compression amplification. The additional acoustic parameters measured were temporal and amplitude parameters. Temporal contrasts between the same speech syllable amplified with linear and nonlinear processing were quantified using the Envelope Difference Index (Fortune et al., 1994). Amplitude measures of speech processed with the different forms of amplification were taken using cone kernel Time Frequency Representation.

In the present study, the initial acoustic analysis data was collated and used to select two contrasting compression amplification conditions for the perceptual experiment. It was hypothesized that the compression condition that best preserved the acoustic parameters of the speech signal (i.e., that was closest to linear) would yield perceptual results similar to linear amplification, that is, the best speech perception results. Conversely, it was hypothesized that the compression condition that most altered the acoustic properties of the speech signal (i.e., that was most different from linear) would yield significantly poorer speech perception results. Thus, no a priori decisions were made about whether the nature of the changes in acoustic characteristics would be good or bad for speech perception but, rather, that the extent of change from linear would be the crucial factor.

In addition to examining the acoustic changes to the speech signal and their perceptual effects, a secondary aim of this study was to investigate the influence of individual listener characteristics on speech perception results with compression amplification. Many studies of speech perception with compression amplification have commented on the high intersubject variability (Benson et al., 1992; Dillon, 1996; Sammeth et al., 1996), variability that has not been fully explained by the subjects' pure-tone audiometric characteristics (Crain and Yund, 1995; Hickson et al., 1995). The individual characteristics included in our research were age, degree of loss, dynamic range, temporal resolution, and frequency selectivity. These characteristics were chosen since they have been shown to explain some additional variance in the unaided speech perception of listeners with hearing impairment, variance that is not accounted for simply by the pure-tone audiogram (Chung and Smith, 1980; Florentine et al., 1980; Patterson et al., 1982; Zwicker and Schorn, 1982; Tyler et al., 1983; Glasberg et al., 1984; Tyler et al., 1984; Schorn and Zwicker, 1990; Florentine et al., 1993; Sommers and Humes, 1993).

METHOD

Participants

Seventy-four adults (49 males and 25 females) were recruited from local audiology services and from the general community. Selection criteria were that participants were aged between 60 and 80 years, had English as their first language, and were able to read and follow written instructions. The mean age was 74 years (SD = 4.3, range = 62–80 years). The audiological criteria were that at least one ear had mild to moderate sensorineural hearing loss, normal middle-ear function, and active acoustic reflexes at levels consistent with cochlear impairment. Table 1 summarizes the pure-tone threshold data for the test ear of participants. The audiometric pattern of sloping mild to moderate sensorineural hearing impairment was chosen as it is most representative of hearing loss in older people and because the amplification provided by the experimental hearing aid was appropriate for this configuration.

| Table 1 | Summary of Hearing Threshold Level Results (dB HTL) for the 74 Participants |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|         | PTAvg. *        | 250 Hz          | 500 Hz          | 1000 Hz         | 2000 Hz         | 4000 Hz         |
| Mean    | 47.12           | 24.79           | 24.18           | 29.26           | 46.55           | 65.54           |
| SD      | 8.64            | 10.12           | 10.72           | 12.71           | 11.73           | 10.87           |
| Range   | 25–73           | 0–50            | 10–50           | 5–60            | 20–70           | 45–95           |

*Pure-tone average at 1000, 2000, and 4000 Hz
Although the vast majority of the participant group (91%) had previously been fitted with hearing aids (37% binaurally, 54% monaurally), only 20% said that they used hearing aids regularly. Sixteen percent reported never using their hearing aids, and 61% said they used aids occasionally at home. Usage rates of hearing aids are somewhat lower than typically reported for older Australians (Dillon et al, 1999; Hickson et al, 1999b), and it is possible that those who responded to advertisements about a research project testing hearing aid systems were more likely to be people who were less satisfied with their aids and wore them less often than other people with hearing impairment.

Materials

Temporal Resolution and Frequency Selectivity Stimuli

Certain characteristics of the psychoacoustic materials used were common to both frequency and temporal measures. For both measures, masked threshold sound pressure levels (SPLs) of a 2 kHz tone were obtained. The tone had a nominal duration of 500 msec inclusive of a 50 msec onset and offset, shaped by a half-hanning function, and was pulsed at a repetition rate of one per second. Masking noises used for both temporal resolution and frequency selectivity tasks were generated digitally using MATLAB. They were 30 sec segments spliced together to produce a temporally continuous signal loop. The sample rate used throughout this procedure and in the digital-to-analogue conversion during data collection was 20 kHz with 16-bit resolution.

Temporal Resolution

A simplified masking period pattern method developed by Zwicker and Schorn (1982), and described by Robinson et al (1984), was used to obtain the measure of temporal resolution from each participant's test ear. In this procedure, monaural threshold SPLs of a 2 kHz tone were determined in three conditions: (1) in quiet (TQ), (2) when masked by a continuous random noise one octave wide and centered on 2 kHz (TC), and (3) with TC modulated by a 14 Hz square wave (TM). These threshold values were used to calculate temporal resolution (TR) as follows:

\[ TR = \frac{(TM - TC)}{(TM - TQ)} \]

This is Zwicker and Schorn's (1982) temporal resolution factor but with the sign reversed (Robinson et al, 1984). With good temporal resolution, TM would normally be better than TC since the listener is able to take advantage of the noiseless temporal gap introduced by modulation in order to better detect the tone. Impaired temporal resolution reduces this advantage so that poor temporal resolution is reflected as an increase in TR in the above equation (to a smaller negative value). Zwicker and Schorn (1982) reported a normal range of TR to be between -1.3 and -0.5.

The octave-band noise was generated by filtering random noise with an elliptical bandpass filter that had attenuation of more than 60 dB/octave in the stop band and whose 3 dB points were one octave wide centered on 2 kHz. The modulated noise was generated by multiplying the octave-band noise with a 14 Hz square wave. The square wave had a 50 percent duty cycle. The result was that the modulated noise had an SPL exactly 3 dB lower than the continuous octave-band noise but an equivalent spectral density. Masker levels were set individually so that one-third-octave-band SPLs in the band centered on 2 kHz were 43 dB above TQ. The exception was this: whenever TQ exceeded 50 dB SPL, a sensation level of 33 dB was used to avoid excessively loud stimuli.

Frequency Selectivity

A measure of frequency selectivity was determined using a notched-noise-masking technique, the detailed principle of which is described fully elsewhere (Patterson and Nimmo-Smith, 1980; Patterson et al, 1982; Robinson et al, 1984). Briefly, though, this method of estimating the selectivity of the auditory filter requires the measurement of two masked thresholds, the first (TB) being the threshold of a 2 kHz tone masked by a broad band noise of uniform spectral density, and the second (TN) the threshold of a 2 kHz tone masked by a noise having the identical spectral density as TB but with a spectral notch centered at 2 kHz. The difference
between these two masked threshold levels (TB–TN) gives the measure of frequency selectivity. The index increases with increased impairment.

The notched-noise masker was produced by placing two noise bands with steep skirts and flat tops symmetrically around the 2 kHz tone on a linear frequency scale. The bands were produced by first low-pass filtering two independently seeded white noise signals to produce two identical noise bands whose frequency ranged from 0–700 Hz. These noise bands were multiplied by an 800 Hz and a 3400 Hz sine wave respectively and then summed together to form the complete notched-noise masker with the notch centered on 2 kHz \( f_n \). The notch width \( \sqrt{f} \) was 0.35, where \( \Delta f \) is the frequency separation between the signal \( f \) and the edge of the notch (in either direction). All measurements of the width of the notch were made at the 3 dB points on the edges of the noise bands; the spectrum level of the noise in the pass band was 40 dB/Hz SPL. The broad band noise was produced from differently seeded white noise samples and was identical to the notched noise in every other respect except that it was continuous, that is, no notch was inserted in the spectrum. It had an identical spectrum level of 40 dB/Hz in the pass band.

**Speech Stimuli**

The speech stimuli used in this experiment consisted of 24 consonant-vowel (CV) syllables. These were pairings of eight voiceless consonants (\( /p/, /t/, /k/, /\beta/, /\theta/, /s/, \) and \( /\ell/ \)) with the three vowels /a/, /i/, and /u/. A detailed account of the initial recording procedure and the procedure used in creating the hearing aid processed speech can be found elsewhere (Hickson et al., 1999a). In summary, the syllables were initially digitally recorded from an Australian male speaker and stored directly onto computer disk using a Computerized Speech Lab (Kay Elemetrics Corporation, Model 4300B) with a sample rate of 20 kHz and 16-bit resolution. The hearing aid processed speech was produced by recording each of these previously stored syllables through a ReSound BT4 hearing aid.

The original hearing aid processed speech was recorded from the ReSound hearing aid set in 13 different conditions with two different input presentation levels (60 and 75 dB SPL) (Hickson et al., 1999a). One of the amplification conditions was linear in both channels, and in the remaining 12 conditions, the compression ratios were either 1, 1.5, or 3 in one or both of the aids' channels, and the crossover frequency varied between .6 and 2.1 kHz. In the perceptual experiment described here, the stimuli used were from the recordings of three of the hearing aid conditions only. These conditions were chosen as a result of the detailed acoustic analysis of the amplified speech signal. Acoustic analysis was conducted using the Computerized Speech Lab (Model 4300, Kay Elemetrics Corporation) and the Computerized Speech Research Environment (CSRE45 [Version 4.5, Avaz Innovations, 1995]). The following acoustic parameters were measured:

1. **Consonant-Vowel ratio (CVR).** Initially, consonant and vowel boundaries were determined by careful examination of the waveform and the spectrogram (including Linear Predictive Coding formant history) and by listening to speech via headphones. The rms (root-mean-square) level of the consonant and vowel portions of the syllable were computed as the square root of the mean of the squared amplitudes (average power) of the sampled points within the time segment (marked by the measured consonant and vowel boundaries) of each sound. The rms level was then converted to dB SPL by taking 20 times the \( \log_{10} \). The rms levels for stops were calculated from the burst and aspiration times, that is, closure was not included (Freyman and Nerbonne, 1989). The CVR for each syllable was calculated as consonant dB SPL minus vowel dB SPL.

2. **Envelope Difference Index (EDI).** Compression amplification has been thought to particularly affect the temporal characteristics of speech due to electronic time delays necessarily introduced by the compression circuitry. The EDI is a technique that allows us to contrast the temporal effects arising from different device settings, such as varying compression ratios and cut off frequencies by comparing their envelopes. The signals to be compared were first digitized at
a sample rate of 20 kHz. Since calculating the EDI between any two envelopes involves subtracting sample point from sample point, the length of the digitized signals were equalized by zero padding the sample point length of the shorter of the two. In order to obtain the signal envelope, a hilbert transform was first applied to each digitized waveform. This technique returns the complex envelope of the signal, the real component of which preserves frequency-related amplitude fluctuations in the final envelope more accurately than by simply taking its absolute value as described by Fortune et al (1994). In order to smooth the envelope further, an overlapping averaging window of +/- 100 sample points was applied. Finally, each envelope was scaled to a mean value of 1.00. Each syllable’s envelope was compared in turn with that syllable recorded in the linear condition (condition 1). For example, the envelope of the syllable /pa/ recorded in condition 1 was compared with /pa/ recorded in conditions 2-13. Envelopes were compared using the formula

\[ EDI = \sum \frac{\text{abs}(env1 - env2)}{2N} \]

described by Fortune et al (1994). Perfect temporal correlation between two signals gives an EDI of 0, and 1.00 is no correlation at all.

(3) Amplitude measures. Three measurements were taken at the same temporal points in each consonant (i.e., onset + 10 msec; midpoint; offset - 10 msec). The cone kernel Time Frequency Representation was used to extract these measures from the spectrograms. This technique was chosen since the resulting representation improves frequency resolution by enhancing spectral peaks in the signal without sacrificing temporal resolution of the signal. A 12 msec analysis window was used with 60 percent overlap and no pre-emphasis.

The difference between condition 1 (linear) and all other 12 conditions was calculated for each of the acoustic parameters (i.e., CVR, EDI, and three amplitude measures) in order to determine the compression amplification conditions that were most similar and most different from linear. For each acoustic value for each consonant, an absolute difference was calculated and converted to a z score. These z scores were then summed across all acoustic characteristics. The amplification condition with the smallest sum was the most similar to linear, and, conversely, the condition with the largest sum was the most different.

The compression ratio and crossover frequency settings chosen for each test condition in the current study are given in Table 2. In the linear condition, both channels were set to provide linear amplification. Compression 1 was the combination of compression settings that produced acoustic characteristics most like the linear condition, and compression 2 was the combination of compression settings that produced acoustic characteristics most unlike the linear condition. Similar acoustic results were obtained for both the 60 and the 75 dB input

![Figure 1](image-url)  
Figure 1 Frequency response of the ReSound BT4 in the three different amplification conditions selected for the perceptual study.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Compression Settings for the Three Test Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Compression Ratio Channel 1</td>
</tr>
<tr>
<td>Linear</td>
<td>1</td>
</tr>
<tr>
<td>Compression 1</td>
<td>3</td>
</tr>
<tr>
<td>Compression 2</td>
<td>1.5</td>
</tr>
</tbody>
</table>

419
levels, and a decision was made to use the stimuli created with the 60 dB input for the perceptual experiment. The frequency response obtained with the three amplification conditions for a 60 dB input level is shown in Figure 1. The frequency response was the same for all participants and, although not customized for individual audiometric patterns, was generally suitable for sloping mild to moderate sensorineural hearing impairment.

Procedure

A brief case history was taken at the beginning of the assessment session. This was followed by routine pure-tone and immittance audiometry assessment to determine suitability of the participant for inclusion in the study. Pure-tone thresholds at octave frequencies from 250 to 4000 Hz were measured using the Hughson-Westlake procedure and an Interacoustics AC4 clinical audiometer calibrated in accordance with Australian Standard 1591.2 (1987). The Amplaid 770 impedance meter was used to measure middle-ear function and to assess acoustic reflex thresholds at octave frequencies between 500 and 4000 Hz. In addition, during this initial part of the session, the participant’s Most Comfortable Listening level (MCL) for speech stimuli was determined adaptively using the Australian version of Boothroyd’s phonetically balanced CVC word lists (AB Words, Boothroyd, 1968). Subsequent assessments were conducted in the test ear only, in the order outlined below.

LGOb (Loudness Growth in Octave Band) Test

Participants’ perception of the loudness of bursts of octave-band noise at various frequencies (500, 1000, 2000, and 4000 Hz) and SPLs was measured using the Resound P3 LGOb system. Each participant heard a series of three bursts of the octave-band noise at each frequency and at each presentation level (20, 40, 60, 80, and 100 dB SPL). After the series, they were asked to indicate by pressing the appropriate button on a control panel whether the sound was (6) too loud, (5) very loud, (4) loud, (3) comfortable, (2) soft, (1) very soft, or (0) cannot hear. The instructions and procedure used for the test were those specified in the ReSound manual (1996; Allen et al, 1990).

Temporal Resolution and Frequency Selectivity

The frequency selectivity and temporal resolution test stimuli were presented using a PC running the CSRE45 software that controlled the presentation of all test stimuli and collected participants’ responses. The masking noises used were previously recorded onto cassette tape and were presented via a good quality cassette player (Yamaha KX330). This output was band-pass filtered (low-pass = 4200 Hz, high-pass = 20Hz) by a Krohn-Hite 3901 filter with skirts of 48 dB/Octave. The test stimuli and masking noise were then independently attenuated by a two-channel computer controlled digital attenuator (Tucker-Davis). Both outputs were then mixed into a single channel (Digitour micro mixer, A-8400) that was then amplified (Marantz Integrated Stereo Amplifier, PM-32) and presented monaurally to the test ear via a pair of TDH49 earphones. Both stimulus and noise masker SPLs produced by the set up were calibrated using a precision sound level meter (Bruel and Kjaer type 1624) attached to an artificial ear (Bruel and Kjaer type 4152) with a one-inch microphone (Bruel and Kjaer type 4176). The pure-tone stimulus SPL was calibrated to Australian Standard 1591.2 (1987). The masker SPLs were calibrated in order to achieve the sound pressure spectrum levels required for each masking noise. A second precision sound level meter (Bruel and Kjaer type 2235) was permanently attached to the amplifier output in order to continuously monitor the level of the electrical signal and to check the system calibration before each participant was tested.

The procedure for measuring the masked threshold of the 2 kHz probe tone was common to both temporal and frequency tasks. Masker and tone were presented to the test ear. Participants indicated that they detected the presence of the tone by clicking a computer mouse on a target placed on the visual display unit before them. Threshold was measured by an adaptive procedure that determined the threshold SPL that corresponded to the 71 percent point on the psychometric function (Levitt, 1971). The initial step size was 8 dB, and the final step
size was 2 dB. The mean value of the last four reversals was accepted as threshold.

**Consonant Perception**

Participants listened to the previously recorded CVs that had been processed by the ReSound hearing aid set in a linear condition and in the two compression conditions described in Table 2. AB word list stimuli were used to determine two presentation levels of the speech stimuli: MCL and MCL-15. The softer presentation level was included in case the MCL level was not sufficiently sensitive to differences across amplification conditions. The first 32 participants were tested at MCL first followed by the softer level, and the order was reversed for the remaining participants. MCLs for the participants ranged from 75 to 105 dB SPL with a mean level of 87.34 (SD = 6.25).

The CVs were presented within the carrier phrase "can you show me <target CV>". Before each amplification condition, participants were familiarized with all CVs by listening to a single unprocessed example of each CV presented within the carrier phrase. No response was required during the familiarization stage. The familiarization material was then repeated as a practice session, but in this phase participants were asked to respond by clicking the mouse in the box on the computer screen containing the correct CV. Feedback was given by lighting the selected box in green for a correct response and lighting the correct box in red for an incorrect response. Participants were required to score above chance in this practice session before moving on to the experimental session.

For each of the amplification conditions, the participant was presented with three blocks of stimuli. Each block comprised the eight consonants /p/, /h/, /k/, /f/, /θ/, /s/, /ʃ/ paired with one of the three vowels /a/, /i/, /a/ four times. The presentation sequence of the three amplification conditions was different for each participant and determined by a Latin Square design. The order of CV presentations within each block and of the blocks within each amplification condition was randomized by the computer software.

**RESULTS**

**Consonant Perception**

Figure 2 shows the participants’ percentage correct consonant identification scores for each of the three vowel categories /a/, /i/, and /a/ at the two presentation levels (MCL and MCL-15dB), in each of the three amplification conditions. Statistical analysis was conducted using a repeated measures 3 x 2 x 3 ANOVA. There were significant main effects of level of presentation (F = 7.45, p < .01) such that higher scores were obtained at MCL than at MCL-15. There were also significant main effects of amplification condition (F = 31.92, p < .001) and vowel category (F = 88.54, p < .001) indicating that perception differed significantly across the amplification conditions and that the vowel context in which the consonant occurred was also significant. The pattern of results for the different vowels was very similar at MCL and at MCL-15, and no interaction effect was found. The interaction between amplification condition and the vowel category was not significant, indicating that the pattern of results for a particular type of amplification did not change for the different vowel environments.

Multiple comparisons of the simple effects associated with listener’s performance in each vowel category were made with Bonferroni adjustments. Consonant perception with /a/ was not significantly different to that with /i/, however, performance with /a/ was significantly better (F = 106.77, p < 0.001) by more than 13 percent than performances in both /i/ and /a/ categories.
There was a significant interaction between amplification condition and level of presentation (F = 11.72, p < 0.001). Post hoc paired sample t-tests were undertaken to investigate the differences between amplification conditions at the two presentation levels. The Bonferroni t-test results with a familywise significance level of p < .05 showed that, at MCL, all amplification conditions were different from each other. The highest mean scores were obtained for the linear condition (68.07%), followed by Compression 1 (65.73%), and then Compression 2 (60.92%). When the same procedure was applied to the results from MCL-15, the only significant difference was that the mean for Compression 1 (62.5%) was different from the mean for Compression 2 (60.07%). The results for the linear condition (61.39%) were not significantly different from the two types of compression.

Relationship between Changes in Consonant Perception and Participant Variables

A secondary aim of this study was to investigate the relationship between differences in performance observed in the various amplification conditions and the participant variables of age, degree of hearing loss, dynamic range, temporal resolution, and frequency selectivity. Change in consonant perception was expressed as a difference score calculated by subtracting the percent correct in the linear condition from the percent correct in the Compression 2 condition for /l/, /l/, and /l/ categories at each level (MCL and at MCL-15 dB). Degree of hearing loss was taken as the test ear pure-tone average at 1000, 2000, and 4000 Hz (see Table 1). Dynamic range at each frequency was defined as the difference (in dB) between the level categorized as (1) very soft, and the point rated as either (4) loud or (5) very loud using the LGOB rating scale. The reason for the variable upper level is that a number of participants did not use the full range of descriptors. Mean dynamic range values were approximately 50 dB at 500 and 1000 Hz and reduced to approximately 30 dB at 4000 Hz, where participants had the greatest hearing loss (see Figure 3). Since there was little difference between the dynamic range measured at 500 Hz and 1000 Hz, these variables were averaged to form a single index of low-frequency dynamic range (Mean = 51.14 dB SPL, SD = 13.73, Range = 19.5–78.5). The 1000 Hz and 4000 Hz dynamic range variables were combined separately to form an index of high-frequency dynamic range (Mean = 35.93 dB SPL, SD = 11.08, Range = 11–64.5). The mean temporal resolution value was -63 (SD = .52, Range = -2.09–13) and the mean frequency selectivity value was -8.97 (SD = 4.62, Range = -20.33–2.44).

### Table 3  Comparison between Percent Correct Identification Scores with Linear and Compression 2 Amplification

<table>
<thead>
<tr>
<th>Level</th>
<th>Vowel</th>
<th>Same Score (n)</th>
<th>Best with Linear (n)</th>
<th>Range of Improvement (%)</th>
<th>Best with Compression 2 (n)</th>
<th>Range of Improvement (%)</th>
<th>$X^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCL</td>
<td>/l/</td>
<td>8</td>
<td>40</td>
<td>3–37</td>
<td>26</td>
<td>3–25</td>
<td>3.0*</td>
</tr>
<tr>
<td></td>
<td>/l/</td>
<td>7</td>
<td>54</td>
<td>3–44</td>
<td>13</td>
<td>3–25</td>
<td>25.1**</td>
</tr>
<tr>
<td></td>
<td>/l/</td>
<td>13</td>
<td>48</td>
<td>3–37</td>
<td>13</td>
<td>3–13</td>
<td>20.1**</td>
</tr>
<tr>
<td>MCL-15</td>
<td>/l/</td>
<td>7</td>
<td>41</td>
<td>3–25</td>
<td>26</td>
<td>3–34</td>
<td>3.4*</td>
</tr>
<tr>
<td></td>
<td>/l/</td>
<td>8</td>
<td>40</td>
<td>1–28</td>
<td>26</td>
<td>3–41</td>
<td>3.0*</td>
</tr>
<tr>
<td></td>
<td>/l/</td>
<td>9</td>
<td>32</td>
<td>0–25</td>
<td>33</td>
<td>0.1–25</td>
<td>0.15</td>
</tr>
</tbody>
</table>

*p < 0.05; **p < 0.001
Bivariate correlation analysis revealed no significant correlations between the difference scores (Linear—Compression 2) and age, degree of hearing loss, dynamic range, temporal resolution, and frequency selectivity. Frequency selectivity and temporal resolution measures were highly correlated with each other ($r = 0.48$, $p < 0.001$) as were the two dynamic range variables ($r = 0.792$, $p < 0.001$). Degree of hearing loss was correlated with both low-frequency and high-frequency dynamic range ($r = 0.25$, $p < 0.05$; $r = 0.30$, $p < 0.001$) as was age ($r = 0.24$, $p < 0.05$; $r = 0.30$, $p < 0.001$). No other significant correlations were observed.

From the difference scores, the proportion of participants who scored higher, lower, or the same in the linear condition compared to Compression 2 was calculated. Table 3 shows that for all vowel and amplification conditions (except for /u/ at MCL-15), significantly more participants performed better with linear amplification than with Compression 2 amplification.

**DISCUSSION**

Overall, the results obtained in this study support the hypothesis that compression amplification that best preserves the acoustic properties of speech relative to linear amplification will yield the best speech perception results. This effect was most pronounced for consonant perception data obtained at participants' MCL, where significant differences were obtained between all three types of amplification. At that level, consonant perception was best with linear amplification, followed by Compression 1 (the condition with least disturbance to acoustic parameters), followed by Compression 2 (the condition that most altered the acoustic properties of the speech signal). Compression 1 had a compression ratio of 3 in the low-frequency channel, and a compression ratio of 1.5 in the high-frequency channel, with a crossover frequency of 1000 Hz. Compression 2 had most compression in the high-frequency channel (ratio = 3), a compression ratio of 1.5 in the low-frequency channel, and a crossover frequency of 2000 Hz. At 15 dB SPL below MCL, the difference was less pronounced, with linear being the same as both compression conditions, but Compression 1 scores were still significantly better than Compression 2. It has been argued that the benefits of compression for speech perception should be evident at softer levels (Humes et al, 1999), and it is true that there was less difference between linear and compression conditions at MCL-15 in the present study. Nevertheless, results with compression were not significantly better than linear at this softer level.

Because both compression ratio and crossover frequency differed in the two compression conditions, it is not possible to determine which had the greater impact. Previous studies have indicated that increasing compression ratio decreases speech perception in noise, but not in quiet (Boike and Souza, 2000; Hornsby and Ricketts, 2001). Crossover frequency effects have not been reported. On the basis of the results of the present study, it can be stated that compression amplification that had an emphasis on compression in the high-frequency channel (i.e., high compression ratio and high crossover frequency) had an adverse effect on consonant perception for older people with sensorineural hearing impairment. A possible explanation for this is that voiceless consonant perception is largely dependent on high-frequency amplification, and compression that focuses on the high frequencies is therefore particularly detrimental. To date, there have been conflicting reports about the effects of high-frequency compression on speech perception. Souza and Turner (1999) tested consonant recognition in 16 people with mild to severe sensorineural hearing impairment. Speech was processed with either linear or two-channel wide dynamic range compression with a compression ratio of 5 in the high-frequency channel, a ratio of 2 in the low-frequency channel, and a crossover frequency of 1500 Hz. At softer input levels (55 and 70 dB SPL), results with compression were better than with linear, and at the highest level (85 dB SPL) there was no difference. Stone et al (1997), on the other hand, did not find any benefits of fast-acting compression in the high-frequency region; in fact, speech intelligibility was significantly worse with compression than with linear amplification. It must be pointed out, however, that this was for speech in noise, rather than in quiet, as was the case in the present study and for Souza and Turner (1999). Another study that provides some
support for the findings here is the work of Bamford et al (1999). They describe positive results with a two-channel hearing aid that has low-frequency compression and high-frequency linear, like Compression 1 in the present study, for a group of 25 children with hearing impairment.

Another incidental finding of the present study was that consonant perception with amplification (both linear and compression) was best in the vowel environment /u/, compared to /a/ and /i/ contexts. This is most probably related to the finding reported by Hickson et al (1999a) that, irrespective of amplification condition, CVRs were highest for syllables containing the vowel /u/.

The second aim of this study was to investigate relationships between differences in consonant perception with linear and compression amplification and participant characteristics of age, degree of hearing loss, dynamic range, temporal resolution, and frequency selectivity. No significant relationships were found, leaving intersubject variability with compression unexplained. Similar results have been obtained previously for the factors of age, degree of hearing loss, and dynamic range (Hickson et al, 1995; Sammeth et al, 1996; van Harten de Bruijn et al, 1997; Souza and Kitch, 2001). The frequency selectivity and temporal resolution data in this study is comparable to that typically obtained from similar populations using similar methods (Patterson et al, 1982; Lutman, 1990). Dreschler (1989) reported a similar finding to the present study in that adults’ temporal resolution was poorly correlated with phoneme perception with both linear and compression amplification. In addition, Franck et al (1999) found no significant difference in adults’ phoneme scores with linear and either single or multichannel compression and found no relationship with notched noise measures of frequency selectivity. Souza and Bishop (2000) did find that audiometric configuration affected outcomes with compression in that listeners with sloping hearing loss showed smaller improvements in consonant recognition than those with flat losses for speech amplified with wide dynamic range compression. A limitation of the present study was that all participants had sloping mild to moderate sensorineural hearing impairment. This was necessary because the same frequency response was used for all participants; however, it meant that it was not possible to investigate the effects of configuration of hearing loss. Another related issue here is that the frequency response would have been more suitable for some participants than others, and the influence of this on the study’s findings is not known.

**CONCLUSION**

The results of the present study indicate a relationship between acoustic changes to the hearing aid processed speech signal and the speech perception performance of older adults with sensorineural hearing loss. It may be that some of the variability evident in research studies with compression amplification highlighted at the beginning of this article may be due to the variable distortion caused by the compression system itself, rather than individual participant attributes. Indeed, no relationship was found in the present study between a number of participant variables and reduced performance with compression amplification. Thus, we agree with Hedrick and Rice (2000), who stress the importance of acoustic analysis of the aided speech signal. It may be possible to define acceptable levels of “distortion” with compression based on acoustic analysis data. It is recognized, however, that speech perception is just one aspect of outcome with compression amplification and that other measures (e.g., self-report of sound quality, satisfaction) are also important. Nevertheless, the majority of people with sensorineural hearing impairment hope to achieve improved speech perception with amplification, and the importance of focusing on perception cannot be overlooked.

**Acknowledgments.** The research was supported by a grant from the National Health and Medical Research Council (Project No.: NHMRC 0012G; Grantor ref.: 971122). The authors would also like to thank Dianne Bates for her assistance with data collection and analysis.

424
REFERENCES


