A COMPARISON OF THE EFFECTS OF COMPRESSION AMPLIFICATION AND DIRECT CONSONANT ENHANCEMENT

by

GIRIJA S. SUNDAR

A dissertation submitted to the Graduate Faculty in Speech and Hearing Sciences
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy, The City University of New York

2004

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

8/26/04 Arlene Neuman
Date

Prof. Arlene Neuman
Chair of Examining Committee

8/26/04
Date

Dr. Martin Gitterman
Executive Officer

Dr. Harry Levitt

Dr. Stanley Gelfand

Dr. Elizabeth Kennedy
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

A Comparison of the Effects of Compression Amplification and Direct Consonant Enhancement

By Girija Sundar

Advisor: Professor Arlene Neuman

Past research has shown that changing the consonant-vowel ratio (CVR) by increasing the consonant level improves speech recognition by persons with hearing loss. Since compression amplification provides an automatic method of increasing CVR, the purpose of this study was to compare the consonant recognition obtained using two-channel syllabic compression to that obtained by increasing consonant level (CE Peak enhancement). A second goal was to examine how the input level to a compression hearing aid would change CV ratio and affect consonant recognition. A comparison of speech recognition scores for the baseline (un-enhanced), compression (65 dB input), and CE Peak conditions revealed no significant difference between the baseline and compression conditions, and small, but significant increases with CE Peak enhancement. An analysis of the CV ratios obtained with compression revealed that often there were large changes in CV ratio characterized by decreases in the level of the vowel in combination with increases in the level of the consonant. The simultaneous change in vowel and consonant levels may be responsible for the poorer recognition with compression than with direct enhancement of only the consonant level. CV ratio increased with increasing input levels. On average, recognition scores increased with increasing input level, but the pattern of performance differed as a function of the consonant.

Acknowledgements

I would like to acknowledge and express my gratitude to the many people who have helped me throughout my academic career. I am especially indebted to my advisor, Professor Arlene Neuman, whose kind patience and vital impatience guided me from the very germination of the idea for my dissertation through its development into a stimulating project. Her rigorous, careful scrutiny, her adherence to scientific methodology, her scrupulous eye for details and her dedication to high standards have been an inspiration. She is a true mentor and I feel honored to have been one of her students.

Distinguished Professor Dr. Harry Levitt served as a member of my committee; his insightful and detailed review of my dissertation---and, indeed, his encouragement throughout my years at CUNY---were invaluable. I wish to thank my friend and committee member Dr. Stan Gelfand for his able assistance and support, and Dr. Elizabeth Kennedy, the external examiner, for her comments. I am particularly grateful to Dr. Arthur Boothroyd, who mentored me during my first few years at CUNY and whose advice, criticisms, and suggestions were absolutely instrumental in my success.

This endeavor would have been impossible without the steady and dependable support of my husband, Kalyan. His sense of humor and perspective kept me grounded; his culinary expertise kept our entire family well fed; and his thoughtfulness and understanding during difficult times made the dissertation process enjoyable. Our children, Arya and Sheila, deserve special thanks. My doctoral pursuit has meant a lot to them, and their unconditional enthusiasm made it easier for me to stray from some of my parental duties.

Four other people deserve specific recognition. Dr. Betty Kollia gave me critical commentary when I was getting lost in the thicket of my data, and her wisdom, encouragement and personal friendship over the years have helped me immensely. Dr. Shantha Madhavan allowed herself to be a ‘sounding board’ while I was analyzing my data; Vinay Krishna navigated me through Microsoft Excel; and my office manager Maryann Wilson held our friendship together in spite of my constantly fluctuating schedules and idiosyncrasies.

I wish to thank Oticon Inc. for generously supplying me with hearing aids for the research project, and Mr. Stephen Crump of Kay Elemetrics Corp., who provided me software for spectral analyses and considerable technical assistance.

Finally and most importantly, I wish to thank two lifelong friends, Raghu Sarma and Rajaram Nityananda, who convinced me to pursue an academic career in the first place. The merits of that career are entirely due to their tireless efforts, and I hope they will think it worthwhile.

TABLE OF CONTENTS

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER IV (Contd.) Page

CHAPTER V

LIST OF TABLES

LIST OF FIGURES

APPENDICES

CHAPTER 1

Introduction

Sensorineural hearing loss results in threshold elevation, reduced dynamic range of hearing, and reduced speech recognition. In general, greater hearing loss is seen in the frequencies above 1000 Hz. In persons with sensorineural hearing loss, the dynamic range of hearing is reduced, i.e., the threshold of hearing is elevated and the threshold of discomfort is reduced, or remains at a level similar to that of people with normal hearing. People with sensorineural hearing loss may also experience loudness recruitment. Loudness recruitment is the rapid growth of loudness above a listener's threshold of hearing as the intensity of sound increases.

For the hearing-impaired person, elevated thresholds result in lack of audibility of some or part of the speech signal. The components of speech sounds are broadly categorized as vowels and consonants. In normal conversational speech, the frequency and the intensity of the components of speech fluctuate rapidly. Vowel sounds are usually more intense than consonants and generally contain more low frequency energy. The difference in intensity between the weakest consonant and strongest vowel can be as much as 40-50 dB (Levitt, 1982). In addition to being weaker in intensity, much of the energy in consonants is in the high frequencies (2,000 to 8,000 Hz). The combination of increased hearing loss at higher frequencies and decreased energy in the speech signal at higher frequencies results in poor audibility and in degraded speech recognition.

Several studies have investigated direct manipulation of the Consonant - Vowel Ratio (CVR) as a method of improving consonant recognition by normal and hearing-impaired listeners (e.g., Hecker, 1974; Gordon-Salant, 1985; Montgomery and Edge,1988; Freyman and Nerbonne, 1989; Kennedy, Levitt, Neuman and Weiss,1998, Smith and Levitt, 1999). The consonant vowel ratio (C-V ratio) is “the ratio of acoustical power in a consonant to that in the adjacent vowel” (Braida et al, 1979). Typically, the CVR of a syllable is calculated by measuring the RMS level of the vowel alone and of the consonant alone. The CVR is the difference between the two levels. Research has shown that manipulation of the consonant - vowel ratio can result in improvements in speech recognition when compared with un-enhanced speech (Hecker, 1974; Gordon-Salant, 1985; Montgomery and Edge,1988; Freyman and Nerbonne, 1989; Kennedy, Levitt, Neuman and Weiss,1998, Smith and Levitt, 1999). What has not been considered in the studies with hearing-impaired listeners is that amplification will automatically change the consonant-vowel ratio. Therefore, in order to determine whether C-V ratio enhancement should be seriously considered as an algorithm to be included in hearing aids, it would be important to compare performance of hearing-impaired listeners with a type of amplification that is likely to modify the C-V ratio to their performance with C-V ratio enhancement.

Syllabic compression is a processing scheme where “the static and dynamic characteristics of the system are chosen to match the temporal level variations that occur within the syllables of speech to the residual dynamic range of the impaired listener” (Braida et al., 1982). Multi-band syllabic compression should increase the CVR by increasing the level of low amplitude consonants and decreasing the level of high amplitude vowels. Several studies have reported that of the various applications of signal processing strategies in hearing aids that improve consonant vowel ratios, syllabic compression holds significant promise (Hickson and Byrne, 1995; Hickson, Dodd and Byrne, 1995; Hickson and Byrne, 1997; Hickson, Thayer and Bates,1999; Preves, Fortune, Woodruff and Newton,1991; Sammeth, Tetzelli and Ochs, 1996).

The principal goal of this experiment was to determine how consonant recognition of voiceless consonants obtained with a two-channel wide dynamic range syllabic compression amplifier compares with optimal consonant recognition obtained using C-V ratio enhancement. The change in both the vowel and consonant energy and the C-V ratio enhancement obtained with compression amplification was also quantified.

CHAPTER 2

Review of the Literature

Several investigators have evaluated direct manipulation of the consonant vowel ratio as a method of enhancing consonant recognition by hearing-impaired listeners. Traditional amplification using compression techniques can also modify the consonant- vowel ratio. Following is a review of research pertaining to the two distinct approaches, namely direct manipulation of consonants to enhance consonant perception and syllabic compression that theoretically can enhance consonant perception.

Hecker (1974) was the first to study the effects of direct manipulation of consonant vowel ratio on speech recognition performance. He hypothesized that the consonant -vowel ratio is related to speaker intelligibility. To test this hypothesis, he manipulated the CVR in recordings of the Modified Rhyme Test recorded by two different talkers. Speech recognition results differed between the two different recordings. In the original recordings, the C-V ratio of the speech recorded by the more intelligible talker was greater than that of the less intelligible talker. In the study, the C-V ratio of the more intelligible talker was reduced to match the C-V ratio of the less intelligible speaker, and the C-V ratio of the less intelligible talker was increased to match the C-V ratio of the more intelligible speaker. The differences in the mean CVR of the two speakers ranged from -4.75 dB for /v/ to + 6.47 dB for /t/. The modified stimuli were presented to 20 normal hearing listeners. Speech recognition scores were obtained for both modified and unmodified stimuli at +4 and –4 signal-to-noise ratios (broadband noise). The increase in the C-V ratio improved word recognition scores by 3.15% at +4 dB signal-noise ratio and 4.28% at -4 dB signal-noise ratio for the less intelligible talker. Decreasing the C-V ratio of the more intelligible talker decreased performance on the word recognition test by 3.6% at +4 dB S/N ratio and 5.73% at -4 dB S/N ratio. Significant improvement (0.05 level) was observed for nine consonant conditions overall. Intelligibility scores for test consonants /t/ and /r/ in the initial position decreased significantly (0.05) when C-V ratio was reduced. Recognition of /t/ and /w/ improved significantly when the C-V ratio was increased. Similarly, intelligibility scores for test consonants in the final position decreased significantly (0.05 level) for /t/, /s/, and /tS/ when C-V ratio was decreased, and improved significantly for /b/ and /l/ when the C-V ratio was increased. This study was the first to examine variations in CVR between talkers. Small, but significant improvements in speech recognition resulted from increasing CVR. Small, but significant decrements in speech recognition resulted from decreasing CVR.

Montgomery and Edge (1988) increased the amplitude of consonants to equal those of the vowels in monosyllabic words. These stimuli were presented at 65 dB SPL and 95 dB SPL to a group of listeners with moderate sensorineural hearing loss. The consonant -vowel ratio enhancement resulted in 10 to 12 percent improvement in speech recognition scores at the lower presentation level, but negligible improvement at the higher presentation level. Montgomery and Edge (1988) suggested that improvement in intelligibility following modification at the lower intensity could be due to increased audibility of the consonants. The lack of improvement at the higher presentation level might be due to these consonants being audible even prior to C-V ratio modification. If this were the case, further increase of the consonant level would have little effect on intelligibility.

Gordon-Salant (1986, 1987) examined the effects of altering C-V ratio on the speech recognition performance of both young and elderly normal hearing (1986) and elderly hearing-impaired listeners (1987). Consonants were increased by 10 dB in a set of nonsense syllables. These stimuli were then presented at 75 dB SPL and 90 dB SPL (6 dB signal-to-noise ratio) in both, normal hearing and hearing-impaired listener groups. In both normal hearing groups, increasing the C-V ratio by 10 dB produced a 16% improvement at the lower presentation level and 11% improvement at the higher presentation level. In those with a mildly sloping loss, a 15% improvement was observed at the lower presentation level compared to a 10% improvement at the higher presentation level. Subjects with a steep loss showed a 16.5% improvement at the lower presentation level and 12% improvement at the higher presentation level. This finding of less improvement at a high presentation level is in agreement with the finding of Montgomery and Edge (1988). These results once again suggest that the gain in recognition is due to increased audibility of the consonant at lower presentation levels after CVR enhancement.

Freyman and Nerbonne (1989) examined the relative contributions of consonant audibility and of CVR enhancement. Eight voiceless Consonant-Vowel syllables (/p/, /t/, /k/, /ts/, /f/, /th/, /s/, and /sh/ in combination with the vowel /a/) were recorded by ten adult male talkers. The investigators modified talker differences by digitally altering consonant-vowel syllables produced by 10 adult male talkers. These stimuli were used to test 50 normal hearing listeners under three different conditions:

(1) C-V ratio was presented in the natural form with the level of the vowel level held constant across all the ten talkers.

(2) The highest C-V ratio among the talkers was identified for each syllable, and the C-V ratio of the other nine talkers was adjusted to equal the highest C-V ratio.

(3) C-V ratio was presented in the natural form with syllables calibrated to consonant levels.

Stimuli in all conditions were presented at 65 dB SPL in the presence of white noise at 0 dB signal-to-noise ratio. The average intelligibility score was 52% when stimuli were presented in their natural form. However, the overall intelligibility score for each talker was not significantly related to the mean C-V ratio and the “most intelligible talker had the lowest overall C-V ratio”. Interestingly, a mean speech recognition score of 63% was obtained when C-V ratio was adjusted to equal the highest C-V ratio amongst the talkers. The highest scores (mean score of 72%) were obtained when consonant levels were adjusted without modifying the C-V ratio. This finding suggests that improved intelligibility was related to absolute consonant levels, rather than to the C-V ratio. Comparing the relative merits of C-V ratio modification and consonant amplitude, this study suggests that the major benefit of CVR enhancement was consonant audibility and this effect was found to be independent of CVR. In addition, differences in intelligibility among talkers could not be explained on the basis of the consonant vowel ratio.

Enhancing consonant-vowel ratio can be done by increasing consonant intensity without altering vowel intensity, or by reducing vowel intensity while leaving the consonant level unaltered. Most studies that have examined the effect of C-V ratio enhancement have focused on the effects of changing consonant levels. Sammeth et al. (1999) evaluated the perceptual effects of changing the C-V ratio by reducing the amplitude of vowels in CV-type nonsense syllables. Consonants /p/, /t/, and /k/ were paired with the vowel /a/. Two normal hearing and six hearing-impaired listeners were tested under four different conditions described below.

Condition1: Each consonant was presented in isolation with the vowel spliced from the CV syllable.

Condition 2: The vowel was added back into the syllable re-creating the original CV syllable.

Condition 3: The vowel level was reduced by 6 dB from the original level of presentation.

Condition 4: The vowel level was reduced by 12 dB from the original level of presentation.

The presentation level was referenced to the /p/ detection level for each listener, and was presented at 20 dB above the /p/ detection level for the consonant in isolation and 14 dB above the /p/ detection level for CV syllables. A P/I (Performance/Intensity) function was obtained in 2-dB-steps from 12 dB below /p/ detection threshold to 12 dB above /p/ detection threshold in the presence of broadband noise at 70 dB SPL and 85 dB SPL. Increasing the intensity of consonants in isolation did not improve identification in the normal hearing or hearing-impaired listeners. Although the addition of vowels contributed to a significant improvement in recognition at all levels, consonant recognition did not improve as a function of decreasing vowel level. This study suggests that enhancement of CVR by reducing vowel amplitude does not contribute to improved intelligibility, possibly because vowel cues are important for speech intelligibility.

Balakrishnan et al. (1996) carried out a study to determine whether consonant intelligibility changes with respect to different C-V ratios and to investigate the relationship between natural CVR and optimum CVR conditions. A second experiment sought to investigate C-V ratio modification effects near threshold, where this condition closely approximates a compression condition. The assumption in the first experiment was that at the supra-threshold condition all speech cues would be audible and that consonant intelligibility would vary with CVR modification, but that the natural CVR would provide the best intelligibility.

Two different groups of 50 normal hearing listeners participated in the experiment. Test stimuli consisted of 22 spectrally smeared VCV (/a-C-a/) stimuli under natural and modified CVR conditions. In order to tease out the effects of spectral differences, and focus on the amplitude differences between the phonemes, the VCV stimuli were degraded spectrally. The spectral cues were "smeared" by equating the levels across the critical bands. The waveform of each stimulus signal was processed through a bank of 21 band pass filters (each filter= 1 critical band), and the gain of each filter was adjusted to equal the combined RMS of the 21 filters for every 5 ms. segment. Thus the gross temporal envelope and the relative amplitude differences within each of the VCV segment were maintained in the natural form, but the spectral content was flattened.

CVR and audibility issues pertaining to voiceless consonants are of particular relevance to the current study. Hence the results discussed in the following section will be limited to seven voiceless consonants (/p/, /t/, /k/, /f/, /th/, /s/, /sh/). Results from the first experimental condition showed that mean consonant recognition scores for the voiceless consonants ranged from 26.5% at –25 dB CVR to 46% at 0 dB CVR, while the mean consonant recognition score for the natural CVR was 45%. Thus, on average, decreasing the CVR decreased scores, but increasing the CVR did not necessarily increase scores.

The maximum score for the stimuli with modified CVR was 78% for /f/ at -20 dB CVR. Interestingly, the natural CVR of -28 dB yielded a recognition score of 80%. The minimum score for the CVR-modified stimuli was 40% for /th/ (–25 dB CVR) while the score for the natural stimulus was 44%. These scores suggest that either CVR modification has little effect on these two consonant conditions or the natural CVR condition is close to the optimum. However, the direction of improvement varied with different consonants. Minimal change was observed for stop consonants, /p/ and /t/. The best score for /p/ was 60% at –10 dB CVR and the best score for /t/ was 60% at –20 dB CVR. Both of these stop consonants reached 58% consonant recognition score at their natural CVR, which was -7.7 dB for /p/ and -15.6 dB for /t/. A linear improvement as a function of increasing CVR was seen for /k/, /s/ and /sh/. The maximum score departed from the score at the natural CVR condition by 20% for /k/, by 34% for /s/ and 40% for /sh/.

In the second experimental condition, where effects of CVR modification were examined at threshold, larger improvement was obtained for the weak fricatives and the largest improvements were observed for the sibilants. Scores also improved with increasing signal-noise ratios. Overall mean scores were poorer at threshold than at the supra-threshold level. The largest mean score for the seven voiceless consonants was 32% at the half compression condition at +7 dB S/N ratio. Scores ranged from 11% in the no compression condition at -3 dB S/N to ratio to 32% at +7 dB S/N ratio. Scores were better under compression conditions compared to no-compression at all three signal-to noise ratios. Maximum benefit was observed for /sh/ at +7 dB S/N ratio with a score of 52%, and for /k/ at +2 dB S/N ratio with a score of 52%.

This study demonstrates that in normal ears, the natural CVR condition is probably optimal for speech recognition for plosives and fricatives, while sibilant recognition allows for the manipulation of a wider range of C-V ratios at supra-threshold levels. For certain classes of consonants, such as the weak fricatives, increasing CVR had a detrimental effect. In supra-threshold conditions, where the weakest of consonants are still above threshold, consonant audibility is not an issue and responses that were affected by CVR modification relate to the C-V ratio per se. On the other hand, responses at threshold suggest that some consonants such as the fricatives may be better identified in the presence of higher CVR (compression condition). C-V ratio did not have any effect on consonant identification for certain consonants such as the stops at supra-threshold levels, but yielded better scores at threshold, indicating that the best CVR might lie somewhere in the middle. Therefore as suggested by the authors, identifying the amount of enhancement that is optimal for consonant recognition for different classes of sounds might have yielded better results. The results obtained at the two different levels of testing suggest the need to compare C-V ratio enhancement to compression techniques in the hearing impaired listeners as well.

Kennedy, Levitt, Neuman and Weiss (1998) evaluated the effects of C-V ratio modification on speech recognition in three groups of hearing–impaired adults. Group 1 had a flat audiometric configuration, Group 2 had sloping audiometric configuration and Group 3 had a steeply sloping sensorineural hearing loss. The purpose of the study was to determine the optimal Consonant-Vowel ratio for seven voiceless and nine voiced consonants (VC format). A performance-intensity function was obtained for levels between Most

Comfortable Listening Level and Uncomfortable Loudness Level in 3 dB steps for each VC. The optimal amount of enhancement for each consonant in each vowel environment was identified from the PI function. Analysis of the pattern of performance with the natural speech revealed that stop consonants had the highest score, followed by strong fricatives, nasal consonants and lastly the weak fricatives. A 35% difference was noted between scores for stops and scores for weak fricatives. Consonant recognition was highest in the context of /a/ followed by /u/ and /i/ in that order, for the un-enhanced condition. Significant improvement in consonant recognition was observed as a result of adjusting consonant-vowel ratio. Overall an average 15% improvement in speech recognition was observed for voiced consonants and 24% improvement was observed for the voiceless consonants.

Consonant recognition as a function of consonant enhancement was maximal for stops and strong fricatives and improvement was greater in listeners with flat and mildly sloping hearing loss configuration. The largest improvements were observed for strong voiceless fricatives (/s/ and /sh/), followed by weaker voiceless fricatives (/f/ and /th/). The next largest improvements were for voiceless stops and the strong voiced fricative /z/. Scores for the nasals showed the least improvement as a function of enhancement. Maximum CVR enhancement with respect to vowel environment was observed in the context of /u/, in both the voiced and voiceless consonants. Mean maximum recognition scores for voiceless consonants were 80% for the flat audiometric configuration group, 85% for the sloping audiometric configuration, and 62% for the precipitous audiometric configuration group.

Smith and Levitt (1999) studied the effect of consonant enhancement on recognition of voiceless and voiced fricatives and stops in children. Voiceless stop /k/ and voiceless fricatives /s/ showed the greatest improvement with enhancement, followed by the alveolar stop /d/. A reduction in consonant recognition as a function of enhancement was observed for the voiced fricative /v/, voiceless labial dental /f/, and voiced labial stop /d/. Younger children demonstrated greater improvement in recognition with enhancement than older children.

Although most studies indicate that enhancing the amplitude of the consonants improves consonant recognition, Ohde and Stevens (1983), Hedrick, Schulte and Jesteadt (1995), and Hedrick and Rice (2000) have pointed out that increasing the amplitude of voiceless stop consonants may result in changes to the cues that signal place of articulation. Ohde and Stevens processed synthetic stimuli to yield a series of CVs where the amplitude of the burst was manipulated over a 20 dB range (from -10 to +10 dB in 5-dB steps) relative to the peak amplitude of the adjacent vowel. The purpose of the study was to determine the specific acoustic properties at consonant release that normal hearing listeners use to identify place of articulation. The F2 and F3 transitions at onset frequencies were varied to correspond to a range of variations in acoustic characteristics appropriate for the identification of /p/ to acoustic properties appropriate for the perception of /t/. The F2 transitions were varied from 800 Hz to 1700 Hz in 100 Hz steps and F3 was varied from 2028 Hz to 2800 Hz in 86 Hz steps. Eleven normal hearing listeners participated in the study. Stimuli were presented at 70 dB above the peak vowel level. Listeners rated the stimuli on a six-point rating scale to indicate if the stimuli sounded more like /p/ or more like /t/. Results showed that as the relative amplitude of the burst in the F4- F5 region increased and became higher than the vowel peak, responses tended to be in the direction of alveolar /t/. When the amplitude of the burst in the F4- F5 region was less than the vowel, the response tended to be more in the direction of /p/. When the starting frequency of F2 was relatively higher than the spectral peak after the release of the burst, responses tended to be more in the direction of /t/. Thus it is evident that perception of voiceless stop consonants may require optimum high frequency amplitude cues as well as optimum transition cues.

Hedrick, Schulte and Jesteadt (1995) used stimuli and methodology similar to those of Ohde and Stevens (1983), with one variation. The burst amplitude was manipulated only at the spectral peak and the stimuli were manipulated from -10 to +10 dB relative to the peak at F4 of the adjacent vowel. The stimuli were presented from 55 dB to 100 dB SPL in 5-dB steps. Twenty-five normal hearing and 5 hearing–impaired listeners participated in the study. One of the goals of this study was to examine whether hearing-impaired listeners can make use of amplitude and transition cues to identify stop consonants. The results showed that normal hearing listeners used the formant transition cues over amplitude cues to make judgments on place of articulation, while the hearing impaired listeners used the relative amplitude of the consonant burst as cue to make correct judgments on place of articulation. Normal hearing listeners tended to show more alveolar responses at high presentation levels, and hearing-impaired listeners showed more alveolar responses at 80 than 60 dB SPL.

Hedrick and Rice (2000) examined perception of voiceless stop-consonants (synthetic CVs) processed through a fast-acting wide dynamic range compression hearing aid in a group of normal hearing and hearing-impaired listeners. The amplitude of the consonant burst was manipulated relative to that of the vowel onset in the F4-F5 region. The F2 and F3 formant transition cues were eliminated in order to remove formant transition cues. The goal of this study was to examine whether normal hearing and hearing-impaired listeners use amplitude cues in order to make judgments regarding place of articulation. The amplitude of the consonant burst was varied to range from -12 dB to +12 dB relative to the peak of the vowel onset amplitude in 3-dB steps. When the amplitude of the consonant burst was greater than the vowel onset in the F4-F5 region, the consonant was more like a /t/, and when the amplitude of the consonant burst was lower than the vowel onset in the F4-F5 region, the consonant was more like a /p/. The stimuli were recorded through a single-channel WDRC circuit that had a 48 dB gain and 2:1 compression ratio and presented at 80 dB SPL to the normal hearing and 90 dB SPL to hearing-impaired listeners. Twenty-five normal hearing and five hearing-impaired listeners were asked to make perceptual judgments as to whether the stimulus was most /p/-like or /t-like. Acoustic analyses were performed, and the amplitude spectra of the hearing aid processed most /p/- like stimulus showed that the amplitude of both consonant and vowel increased with amplification, and more importantly, although the vowel amplitude was higher than that of the burst, the amplitude of the burst increased more than the amplitude of the vowel in the high frequency region. In contrast, for the most /t/-like stimulus, in the unaided condition, the burst amplitude was higher than that of the vowel-onset region and when processed through the hearing aid, the amplitude of the vowel-onset high frequency region increased but the burst amplitude remained higher than the vowel amplitude. The neutral stimulus where the relative amplitude value between /p/ and /t/ was in the middle of the continuum between most /p/-like and most /t/-like, showed that hearing aid processing increased the amplitude of the burst, and made it higher than the vowel amplitude in the high frequency region, making it resemble the /t/-like stimulus. The perceptual results indicated that in general, the listeners had more difficulty categorizing amplified stimuli. Those listeners who were able to categorize the stimuli, labeled the stimuli as the alveolar /t/, rather than the labial /p/.

In the studies of C-V ratio enhancement that have been reviewed, the adjustment to C-V ratio has been carried out primarily by using digital waveform editing techniques. More practical methods of achieving audibility and/or C-V ratio enhancement include frequency-specific linear amplification and wide dynamic range compression amplification. As observed by Dillon (2001), it is not clear whether complex algorithms to enhance consonant amplitude will yield significantly better consonant intelligibility than syllabic compression amplification. Non-linear amplification has the potential to increase C-V ratio, as well as consonant audibility. Theoretically, enhancement of consonants can be achieved using more conventional methods of amplification. Linear amplification with high frequency emphasis can alter CVR by providing more gain in the high frequency region thus making consonants more intense. However excessive high frequency gain is often constrained by reduced tolerance for these signals and poor sound quality, (Villchur, 1982; Dillon, 2001). Compression amplification is becoming the preferred method of ensuring audibility of sounds while maintaining user comfort.

In compression hearing aids, the gain is related to the input level in a non-linear manner. Typically, more gain is provided for low input signals and less gain for higher input signals. Signals are amplified in a linear manner until reaching the compression threshold, the level at which the compressor is activated. The amplifier gain is reduced beyond this point by a specified factor, the compression ratio, in order to ensure that the output of the aid does not exceed discomfort levels. The compression ratio is the ratio of change in input to the corresponding output. For example, if an increase of 30 dB at the input level results in 10 dB increase at the output level, the ratio is 3:1. When signals reach the compression threshold, the time it takes for the compression circuit to reduce the gain to the specified limiting level after it reaches its peak level is referred to as the attack time. The time taken by the circuit to return to a linear gain function is called the release time.

Compression amplifiers in hearing aids may be designed in many different ways and for different purposes. In order to manipulate the CV-ratio within a syllable, a syllabic-type compressor would appear to have the most promise. According to Dillon (Dillon, 1996), syllabic compression is characterized by having relatively short attack and release times, a low compression threshold, and relatively low compression ratios. These parameters, theoretically should increase the amplitude of low intensity weaker sound (consonants) while reducing the gain of the more intense sounds (consonants). They should also preserve the relative amplitude of the different sounds within the hearing impaired person's dynamic range (Dillon, 1996). Speech sounds vary over a wide range of intensities. The consequence of fast-acting compression is that the intensity differences between the more intense and less intense phonemes in speech become smaller, thus increasing the consonant-vowel ratio. A review of the effects of compression amplification on C-V ratio and compression amplification on audibility follows.

Hickson and Byrne (1995) investigated how compression amplification might change the CV-ratio of amplified speech. They compared the acoustic characteristics of nonsense syllables processed through linear amplification and compression conditions (compression ratios set to 1.3 and 1.8). The exact compression threshold was not specified, although the authors mention that the "compression activation threshold" was low. Attack time was set at 6 ms and release time varied from 50-500 ms (adaptive compression circuitry). Twenty-seven nonsense syllables from the CUNY Nonsense Syllable Test (Resnick, Dubno, Hoffnung and Levitt, 1975), consisting of 15 CV and 12 VC syllables, were recorded onto tape by an Australian male speaker. In the CV format, only the vowel /a/ was used. For VC syllables the vowels /a/, /u/ and /i/ were used. These stimuli were played through a loudspeaker and then recorded again through an in-the-ear (ITE) hearing aid fitted to KEMAR. Stimuli were also recorded without a hearing aid through KEMAR. The processed stimuli were recorded onto a digital tape and acoustically analyzed.

Duration, intensity and spectral peaks of the un-amplified and amplified utterances were compared. No significant differences were noted for consonant duration or spectral peaks as a function of linear or compression conditions in fricatives, stops or liquid glides. CVR was the one factor that was altered as a function of the amplification condition. In 15 of the syllables processed through the compression amplifier, the C-V ratio was increased due to a combination of an increase in consonant level and a decrease in vowel level. Overall, both compression conditions yielded an increase in consonant levels. For the compression ratio of 1:1.8, consonant levels were significantly higher ($p < 0.05$) for the syllables /ip/, /up/, /ap/, /ad/, and / ta/. Most syllables showed a trend toward increased CVR as a function of compression amplification, although not all were statistically significant. The overall CVR increase from linear to compression condition ranged from 0 to 6.8 dB, with an average value of 2.65 dB.

Hickson, Thayer and Bates (1999) carried out a more thorough investigation to determine how multiple parameters of compression amplification and input level might affect C-V ratio. Compression ratio (CR) and crossover frequency were varied in a two-channel syllabic compression hearing aid system. Consonant-Vowel syllables were recorded through one linear (CR 1:1) and 12 different compression conditions at 60 and 75 dB SPL input levels. Twenty-four consonant-vowel syllables (eight voiceless consonants paired with the vowels, /a/, /i/, and /u/) were recorded through a two-channel behind-the-ear hearing aid. Compression threshold was kept at 45 dB with fast time constants, (attack time < 1 ms.; release time from 5 to 140 ms.). Acoustic analyses showed that the highest CVRs were in conditions with a high compression ratio in the high frequency channel, and with the crossover frequency set to the higher frequency region. Greater CVR values were observed for the 75 dB input than for the 60 dB input, and all of the different compression ratio conditions had significantly higher CVRs than the linear condition. C-V ratio increases were not always due to increases in consonant level or decreases in vowel level. It is evident from this study that compression amplification alters CVR and this alteration varies as a function of compression ratio, crossover frequency, input level, and vowel environment.

The studies reviewed thus far indicate that manipulation of C-V ratio may result in improved consonant recognition. In most studies, improvements occurred in those situations where audibility of the consonant was improved. The results of the Kennedy et al study (1998) seem to indicate that maximum speech recognition enhancement will occur when the C-V ratio is set individually for each phoneme, in each vowel environment, for each listener. In that study a larger improvement in consonant recognition (15%) was obtained than in other studies of CV-ratio enhancement. The practical issues of implementing individualized C-V ratio parameters in hearing aids can become extremely complex. A more automatic method to obtain C-V ratio enhancement is compression amplification. Hickson et al (1995, 1999) have demonstrated a decided CVR advantage with multi-band syllabic compression amplification. The NAL-NL1 method of prescribing compression parameters (Byrne, et al., 2001) is designed to provide maximum speech intelligibility. The goal of this algorithm is to provide maximum speech intelligibility for any speech input level without exceeding the hearing- impaired listener's tolerance levels. This prescriptive procedure allows for the gain and the compression ratio to be set at each frequency. The gain prescribed for each frequency depends on the threshold at that frequency, the three frequency average threshold (500, 1000 and 2000 Hz), and the slope of the hearing loss from 500 to 2000 Hz.

The present study was designed to answer the following questions:

1. What is the effect of syllabic compression (as prescribed by the NAL-NL1 prescriptive method) on C-V ratio for average conversational speech level (65 dB input)?
2. What is the effect of syllabic compression on recognition of voiceless consonants presented at average conversational speech level?
3. How do the changes in C-V ratio obtained with compression compare to the C-V ratio identified as CE Peak?

4. How does recognition of voiceless consonants processed using compression compare to recognition of the unenhanced speech presented at MCL and to the C-V ratio enhanced stimuli (CE Peak) at MCL?
5. How does changing the speech level at the input to the compression hearing aid change the C-V ratio?
6. How does changing input level to the compression hearing aid affect recognition of voiceless consonants?

CHAPTER 3

Methods

Participants

Twelve hearing-impaired adults participated in the study. All participants were native English speakers. Moderately hearing-impaired listeners with flat to mildly sloping hearing loss were targeted for this study, since it has been demonstrated that persons with these type of hearing loss perform similarly with CVR enhancement (Montgomery and Edge, 1988; Gordon-Salant, 1987; Kennedy, Levitt, Neuman and Weiss, 1998). The group consisted of five females and seven males. The mean age of the group was 57.1 years (range of 25-73 years). Audiometric and personal data are shown in Table 1.

All participants had sensorineural hearing loss of presumed cochlear origin as determined through pure-tone air and bone conduction threshold tests at frequencies from 250 Hz-8000 Hz and acoustic immitance procedures. All tests were performed using ASHA recommended procedures (ASHA, 1979; 1985). Participants exhibited normal (Type A) tympanograms with tympanometric peak pressure between -75 and +75 daPa. Tone decay testing at 2000 Hz did not reveal evidence of retro-cochlear pathology in any of the listeners. Word recognition scores obtained in the test ear using recordings of the Auditec NU-6 word list at most comfortable levels of listening appear in Table 1. Thresholds were found to be no better than 40 dB and no poorer than 60 dB across the audiometric frequencies of each listener.

Table 1. Audiometric and personal data of participants (P).

P	AGE	GENDER	EAR	.5 kHz (dB HL)	1 kHz (dB HL)	2 kHz (dB HL)	3 kHz (dB HL)	4 kHz (dB HL)	WDS
1	65	Female	Right	40	50	50	50	55	80%
2	40	Female	Right	35	40	45	50	45	92%
3	73	Male	Left	30	35	50	45	55	72%
4	56	Female	Right	35	40	50	55	50	80%
5	49	Male	Left	45	55	50	60	55	80%
6	68	Male	Right	40	50	45	55	55	80%
7	73	Male	Right	40	45	50	45	55	84%
8	66	Male	Right	35	45	50	60	55	80%
9	61	Female	Right	40	45	50	50	55	84%
10	25	Male	Right	40	40	50	55	50	92%
11	71	Female	Left	40	50	45	50	50	88%
12	38	Male	Right	50	55	55	55	55	88%
MEAN	57.1			39.17	45.83	49.17	52.5	52.92	83%

Test Stimuli

A full description of the creation of the male recording of the CV-ratio enhanced stimuli appears in Kennedy (1998) and in Kennedy et al. (1998). Briefly, these stimuli were digitized using a 20 kHz sampling rate and 16-bit resolution and stored in a 386-based personal computer. Using the waveform editing process proposed by Dubno and Levitt, (1981), the consonant and vowel portion of each nonsense syllable were segmented and the RMS level was measured. The consonant portion of the syllable was enhanced in 3-dB steps from the levels from the baseline (unprocessed) condition to a maximum of +24 dB. For the current experiment, the Kennedy stimuli were converted to .wav files (sampling rate 22,050, 16 bit resolution).

Only the voiceless consonant subsets of the CUNY Nonsense Syllable Test (NST) used by Kennedy et al (1998) were used for this study in order to permit testing of the large number of conditions in a repeated-measures design. The voiceless subsets were chosen because Kennedy et al. (1998) found greater improvement in performance with CVR enhancement for voiceless than for the voiced subsets of the NST.

The three sets of VCs include seven voiceless consonants in combination with three vowels (/i/,/a/, and /u/):

Set 1. /af/; /ak/; /ap/; /as/; /ash/; /at/; and /ath/

Set 2. /if/; /ik/; /ip/; /is/; /ish/; /it/; and /ith/

Set 3. /uf/; /uk/; /up/; /us/; /ush/; /ut/; and /uth/

The nonsense syllable stimuli were processed using an Oticon Digifocus II compact hearing aid whose compression characteristics were adjusted according to the NAL-NL-1 prescriptive procedure for an audiogram representative of the mid-range of the audiogram representing the listener inclusion criteria. The target gain for the audiogram of each listener did not differ by more than 7 dB from the NAL-NL1 targets at any frequency. Recordings were made at three input levels: 50, 65, and 80 dB SPL.

Calibration of Direct Audio Input to the Hearing Aid

The relationship between the response of the hearing aid to electrical input (via DAI) and to acoustic input (passed through the hearing aid microphone) was verified using the instrumentation shown in Figure 1.

The SOUNDCHECK™ computerized test system was used to measure the frequency response. The output of the soundcard of the computer was sent to the loudspeaker of a Frye Hearing Aid test box. The Digifocus hearing aid, programmed as a linear hearing aid with 0 dB gain, was mounted on a 2cc coupler that was attached to the Bruel and Kjaer Type 2218 Sound Level Meter and a Type 4144 measurement microphone. The frequency response was measured using sweep frequency signals of 50, 65 and 80 dB SPL. The AC output of the sound level meter was sent to the input of the computer for measurement of the frequency response.

Once the frequency response was established for the hearing aid with acoustic input, the equivalent voltage required to yield the same frequency response for an electrical signal to the direct audio input (DAI) of the hearing aid was established. The output of the computer soundcard was introduced into the hearing aid via Direct Audio Input (DAI) through an audio boot (Oticon #401). In other regards, the instrumentation was identical to that used for the measurement of the hearing aid with acoustic input. The stimuli were introduced into the hearing aid through direct audio input (DAI). The attenuator on the input to the DAI boot was adjusted in order to obtain a frequency response equivalent to that obtained with acoustic input at each of the three input levels.

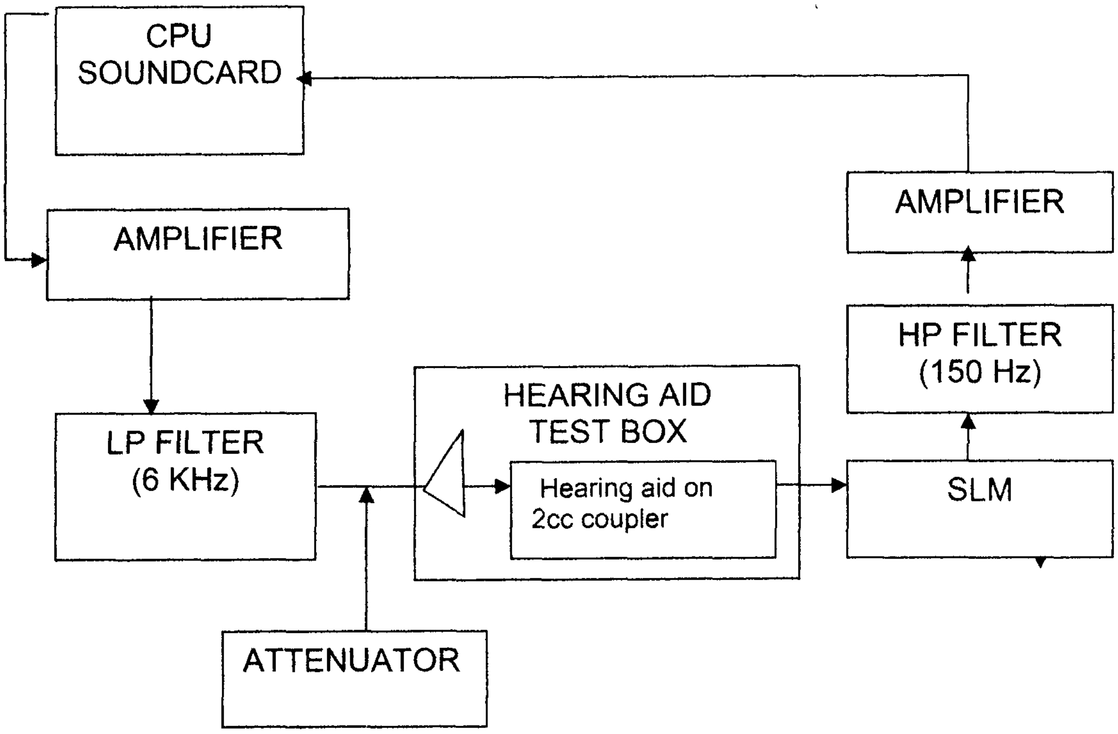


Figure 1. Instrumentation for measurement of the frequency response of the hearing aid with acoustic input or with electrical input (DAI).

The compression characteristics of the hearing aid were verified in the following manner. Input-output curves for both acoustic and electrical input were run at the maximum compression setting for inputs of 250Hz, 750Hz, 1250Hz, 2000Hz, 3000Hz, 4000Hz and 5000Hz, in order to determine the knee point. Comparable knee points were obtained for acoustic and electrical input. An example of the measurement obtained at 750 Hz is shown in Figure 2.

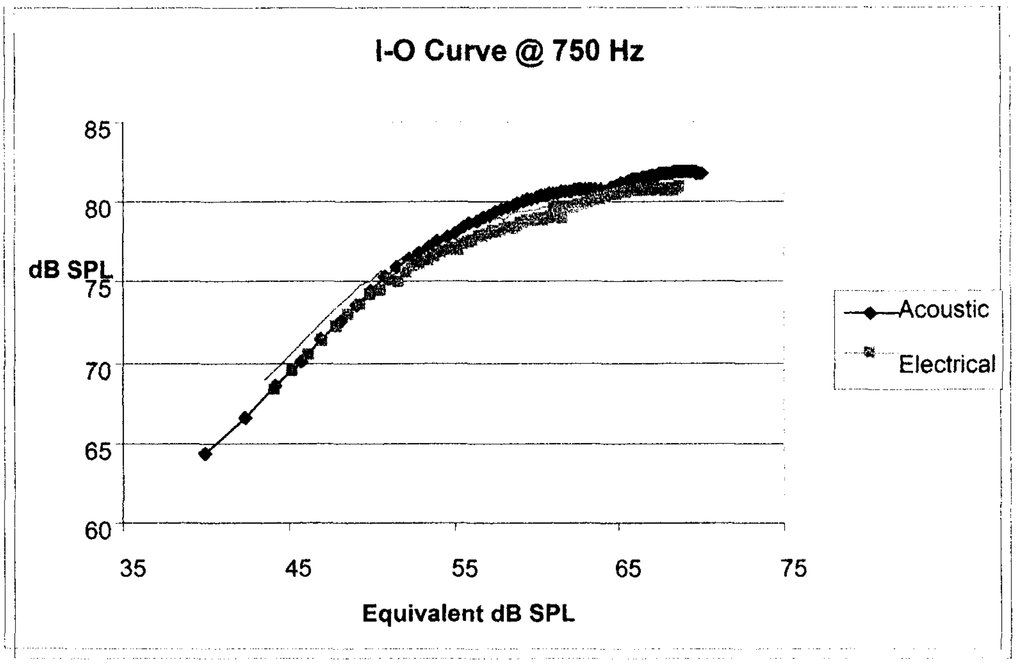


Figure 2. Input-output function of acoustic and electrical input for 750 Hz.

Setting the Compression Parameters of the Digifocus Hearing Aid

The NAL-NL1 prescription was used to establish the gain, compression ratio and compression threshold for the audiogram representing the midrange of the listener audiometric range. The target gains as a function of input level for the mid-range audiogram are listed in Table 2.

The Otiset program was used to set the parameters of the Digifocus hearing aid. The Digifocus is a two-channel hearing aid. The crossover frequency was set at 2500 Hz. The gain at each frequency and the compression ratios were set according to the NAL-NL1 recommendation for each channel. The attack and release times were set to 0 and 40 ms. respectively. These two values represent the fastest attack and release times that the Digifocus II hearing aid allows and fall within the range of values for syllabic compression.

Table 2: NAL-NL1 Target Gains in dB SPL (in 2cc coupler) as a function of input levels (50, 65 and 80 dB SPL).

	500 Hz	1000 Hz	2000 Hz	3000 HZ	4000 Hz
50 dB input	16	24	26	33	32
65 dB input	11	17	20	28	31
80 dB input	4	10	11	20	19

Recording of the Compression Stimuli

All 21 VC stimuli (presented in a carrier phrase) were recorded through the Digifocus II hearing aid coupled to the left ear of a Knowles Electronic Manikin for Electro-Acoustic Research (KEMAR). The stimuli were sent to the Digifocus hearing aid through DAI at each of the three input levels using the same instrumentation as was used for the calibration procedure. An ER-11 ½" microphone was placed at the medial end of the Zwislocki coupler (DB 100) and the output of the ER 11 amplifier was sent to a TEAC DA-P20 Digital audio tape recorder (DAT).

The "You will Mark" carrier phrase was also recorded at the equivalent of a 65 dB level for use in the measurement of MCL levels. A speech-spectrum-shaped noise signal to be used for calibration was also recorded.

After the stimuli were recorded at the three input levels, they were played back in analog form and re-digitized on a PC for ease of testing.

Instrumentation for Data Collection:

All testing was carried out in a sound-treated, single walled IAC test booth. Audiological testing was performed using a Madsen OB822 audiometer. Air conducted signals were presented via ER-3A insert earphones. Bone conducted signals were delivered via a Radioear B-71 vibrator. Speech stimuli for speech audiometric testing were delivered from a DENON compact disc player (DCD-425). The test stimuli were stored on the hard-drive of a COMPAQ ARMADA E 500 laptop computer and presented using a Maestro Wave sound card. The computer output was sent to the audiometer for appropriate attenuation and signals were presented through ER-3A insert phones.

At the beginning of the experiment, an acoustic calibration of the audiometer was done and measurements of the voltages sent to the insert phones were noted (with a 70 dB calibration signal). Thereafter daily calibration consisted of electrical calibration. In addition, a measure of the voltages for each participant's MCL was obtained and checked at the time of each visit.

Each test material had a speech-spectrum shaped noise used to set the gain on the Tape Inputs to the audiometer. The calibration signal was equal in RMS to the RMS of the word 'mark' in the carrier phrase of the test stimuli.

Subjects were tested individually in a sound treated single walled IAC booth. All listeners were tested in three 2 ½ hour sessions.

Stimuli were presented using a closed set procedure via the laptop computer (Compaq Armada). There were twenty replications of each test condition. On each trial, listeners indicated their response using the mouse to select one of the seven choices from the test consonants. During the testing each listener was given a 10 minute rest after each segment of the test. Each segment included all three vowel conditions in each of the hearing aid conditions.

Session 1: Audiometric testing, MCL measure for compressed and consonant enhancement conditions, UCL measures for each CVR syllable. The instructions and procedures used for measuring MCL and UCL may be found in Appendix 1a and1b, respectively.

Sessions 2 and 3: Test order of the three test conditions (Baseline, Consonant Enhancement, and Compression) was randomized. The effect of compression amplification was assessed at three speech input levels, 50, 65, and 80 dB SPL (COMP50, COMP65 and COMP80).

To determine CE Peak, a P/I function was obtained with the consonants enhanced from12 dB to 24 dB. The scores (in proportions) were calculated. The highest score was identified from among the different levels. The test level that gave the highest score was used as the CE Peak condition. The C-V ratio at that level and the test score were used in all of the analyses.

CHAPTER IV

Results

Data obtained in this study include speech recognition results for five experimental conditions as well as acoustic and spectrographic analyses of the test stimuli. The five conditions are sub-divided into two sets to allow for specific comparisons. One set of comparisons is between processing conditions presented at comfortable listening levels. These include the COMP65 (COMP65), Consonant Enhancement Peak (CE Peak) and Baseline (BSL) conditions. While these three experimental conditions place the speech at the listener's comfort level, they differ with regard to the type of processing, thus resulting in different C-V ratios. The output level for these three conditions was the listener's Most Comfortable Listening Level (MCL).

In the second set of comparisons, the effect of compression at different input levels is assessed with regard to speech recognition performance and the concomitant changes in C-V ratio caused by the compression. The output level of the COMP65 condition was at MCL. The output levels resulting from the lower and higher input levels were determined by the parameters of the compression hearing aid used for processing. The COMP50 condition yielded a lower output level than the COMP65 condition. The COMP80 condition yielded a higher output level.

Effect of Signal Processing at Comfortable Listening Levels:

Table 3 illustrates the different levels of presentation for each listener under each condition. Note that the presentation level for COMP65 is usually similar to the Baseline (BSL) and Consonant Enhancement Peak (CE Peak) conditions.

Table 3. Presentation Levels for Each Subject at Comfortable Listening Conditions In dB SPL

	BASELINE (dB SPL)	COMP65 (dB SPL)	CE PEAK (dB SPL)
Subject: 1	111	116	111
Subject: 2	103	106	103
Subject: 3	106	106	106
Subject:4	108	115	108
Subject:5	112	112	112
Subject:6	108	110	108
Subject:7	115	117	115
Subject:8	113	116	113
Subject:9	109	112	109
Subject:10	109	113	109
Subject:11	105	105	105
Subject:12	113	115	113

The Effect of Signal Processing on Consonant-Vowel Ratio:

The hypothesis of this experiment was that compression would alter consonant vowel-ratio, which, in turn, would result in improved consonant recognition comparable to that obtained with the idealized CE Peak condition. Accordingly, C-V ratio measurements were obtained on all of the stimuli to determine the C-V ratio for the unprocessed (baseline) condition and examine how C-V ratio had been modified by the two methods of signal processing. The root-mean-square (RMS) levels of each consonant and vowel were determined using SoundForge software (SONIC FOUNDRY Version 4.5). The RMS levels for stops were measured after the release of the burst. Recall that C-V ratio in the context of this study is defined as the difference between the intensity (in dB) of the consonant of the VC and that of the vowel. This mode of reference is being used in order to be consistent with the definition of consonant-vowel ratio used in the past (Hecker, 1974; Gordon-Salant, 1985; Montgomery & Edge, 1988; Freyman & Nerbonne, 1989; Kennedy, Levitt, Neuman & Weiss, 1998, Smith and Levitt, 1999). The measured consonant-vowel ratios for each syllable in each of the comfortable listening conditions are shown in Table 4. Note that the larger the number, the better the C-V ratio. The most favorable C-V ratio is indicated with an asterisk.

Table 4. C-V Ratio in the comfortable listening conditions

SYLLABLE	BSL	C 65	CE PEAK
ap	-28	-14.1*	-16.59
ip	-22.37	-21.3	-14.16*
up	-24.86	-5.46*	-13.64
at	-24.16	-13.39	-11.91*
it	-15.47	-8.69	-2.47*
ut	-20.93	-2.45*	-9.52
ak	-26.46	-19.64	-14.76*
ik	-21.57	-11.15	-6.04*
uk	-24.58	-9.08*	-14.98
af	-28.42	-22.45	-17.3*
if	-21.57	-21.28	-6.04*
uf	-21.03	-4.21*	-9.88
as	-11.49	-3.86	-1.72*
is	-2.06	-4.81	9.56*
us	-7.68	10.49*	3.57
ash	-10.65	-1.02*	-1.19
ish	-10.15	-6.95	1.64*
ush	-12.11	9.05*	-1.29
ath	-23.52	-21.43	-16.2*
ith	-19.25	-20.95	-8.75*
uth	-23.44	-7.99	-6.59*

*Asterisk: Highest C-V ratio.

Recall that in the CE Peak condition the consonant level was manipulated to yield higher C-V ratios. Hence C-V ratios for all syllables in the CE Peak condition are higher than in the Baseline condition.

In most cases, compression increased the C-V ratio. For eight of the syllables, the C-V ratio for the compressed speech was actually higher than the C-V ratio of the CE Peak condition. The CE Peak condition had the highest C-V ratios in the context of the vowel /i/ for all of the consonants. In the COMP65 condition, on the other hand, the highest C-V ratios were obtained in the context of /u/ for all the consonants, with the exception of /th/. In the /a/ context, the highest C-V ratios were obtained for /p/ and /th/ in the COMP65 condition. It is apparent from this table that compression improved the C-V ratio from the baseline condition for all but two of the syllables (/is/ and /ith/). Figure 3 illustrates the changes in C-V ratio (from baseline) caused by compression. The magnitude of change was the largest for /ush/ (15.45 dB) and negative for /ith/ and /is/ (~-2).

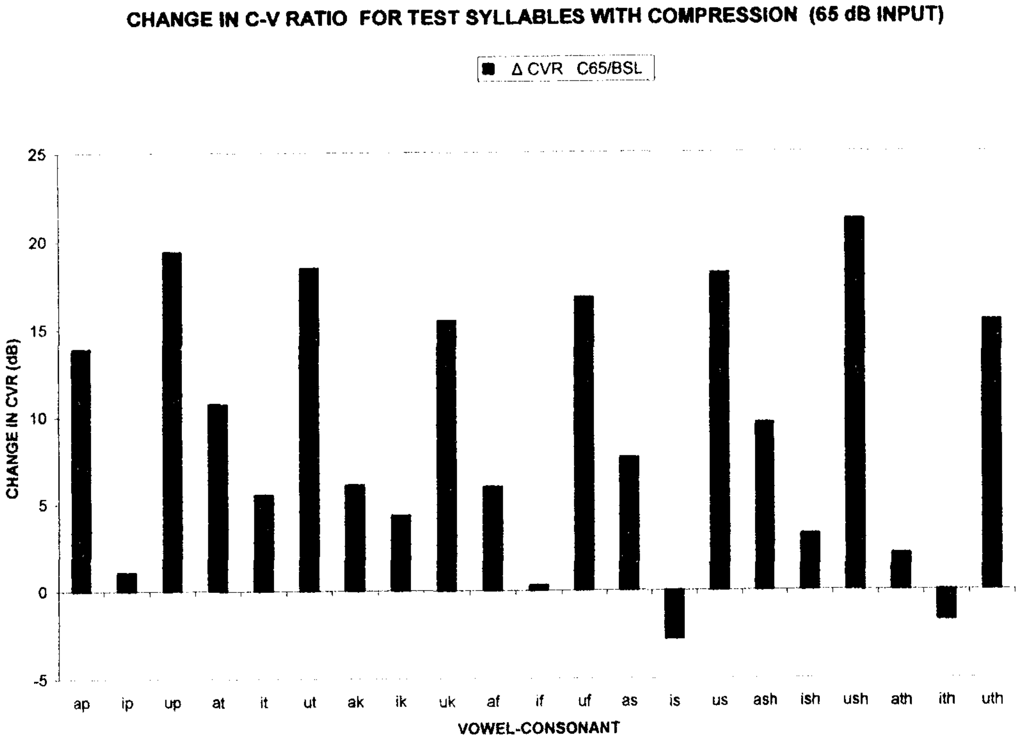


Figure 3. Change in CV-ratio from Baseline for test syllables with Compression (65dB input).

Figure 4 shows the difference in C-V ratio between the COMP65 and CE Peak condition. In eight out of the 21 VC pairs (38%) the compression condition had higher C-V ratios than the CE Peak condition. Higher C-V ratios are seen primarily for syllables containing the vowel /u/.

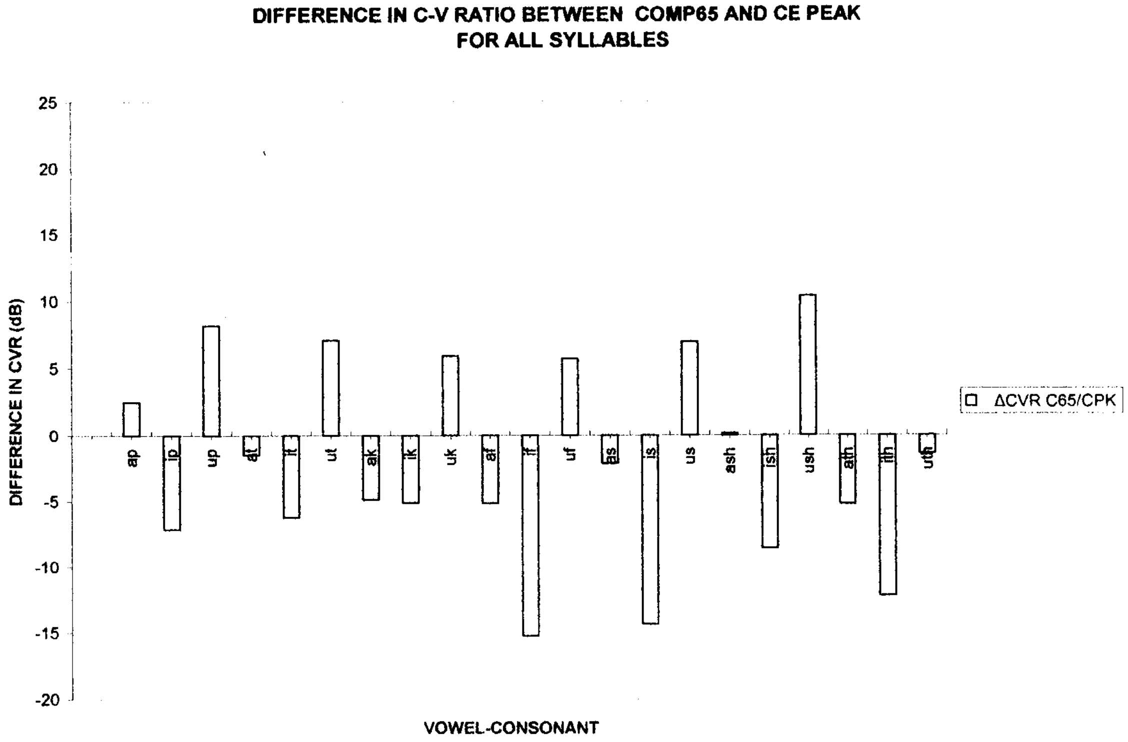


Figure 4. C-V ratio difference between COMP65 and CE Peak for all syllables.

Speech Recognition Performance:

In the first set of comparisons (Baseline, COMP65 and CE Peak), stimuli were presented at most comfortable levels of listening. Speech recognition scores obtained under these three conditions were analyzed in order to allow for direct comparisons among conditions where overall levels are comparable. Data in proportions were submitted to arc-sine transformation prior to analysis. A repeated measures analysis of variance was completed for all conditions, with Hearing Aid (H), Consonant (C), and Vowel (V) conditions as independent variables, and consonant recognition score as the dependent variable. The Hearing Aid conditions were Baseline, COMP65 and CE Peak. The seven voiceless consonants were /p/, /t/, /k/, /f/, /s/, /sh/, and /th/. The vowels were /a/, /u/, and /i/. Statistical significance was set at 0.01. The Post-Hoc analysis of all significant main effects and interactions was performed using the Tukey Honestly Significant Difference (HSD) at the 0.01 level (Winer, 1971). After analysis, arcsine values were converted to percentages for ease in interpretation.

A summary of the results of the ANOVA is shown in Table 5. The main effects of hearing aid condition, consonant, and vowel environment were all found to be highly significant. Many of the interactions were also found to be significant.

Table 5. Summary of Repeated Measures Analysis of Variance for the Comfortable Listening Hearing Aid Conditions.

SOURCE OF VARIATION	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARES	F	SIGNIFICANCE
H	10.10705	2	5.05353	26.313	0.001
V	17.18268	2	8.59134	18.555	0.001
HV	1.12407	4	0.28102	2.785	0.0374
C	211.56664	6	35.2611	22.705	0.001
HC	26.27888	12	2.18991	8.157	0.001
VC	47.749	12	3.97908	10.379	0.001
HVC	10.93183	24	0.45549	2.813	0.001
S	29.58197	11	2.68927		
HS	4.22521	22	0.19206		
VS	10.18636	22	0.46302		
HVS	4.43909	44	0.10089		
CS	102.49785	66	1.553		
HCS	35.43868	132	0.26847		
VCS	50.60479	132	0.38337		
HVCS	42.75251	264	0.16194		
TOTAL	604.66663	755			

The mean consonant recognition scores as a function of Hearing Aid condition are shown in Figure 5. A Tukey post-hoc analysis revealed that the mean score of 91% obtained in the CE Peak condition was significantly higher ($p < 0.01$) than mean performance in the other two hearing aid conditions. The mean scores in the Baseline (85.47%) and COMP65 (81.68%) condition did not differ significantly from each other.

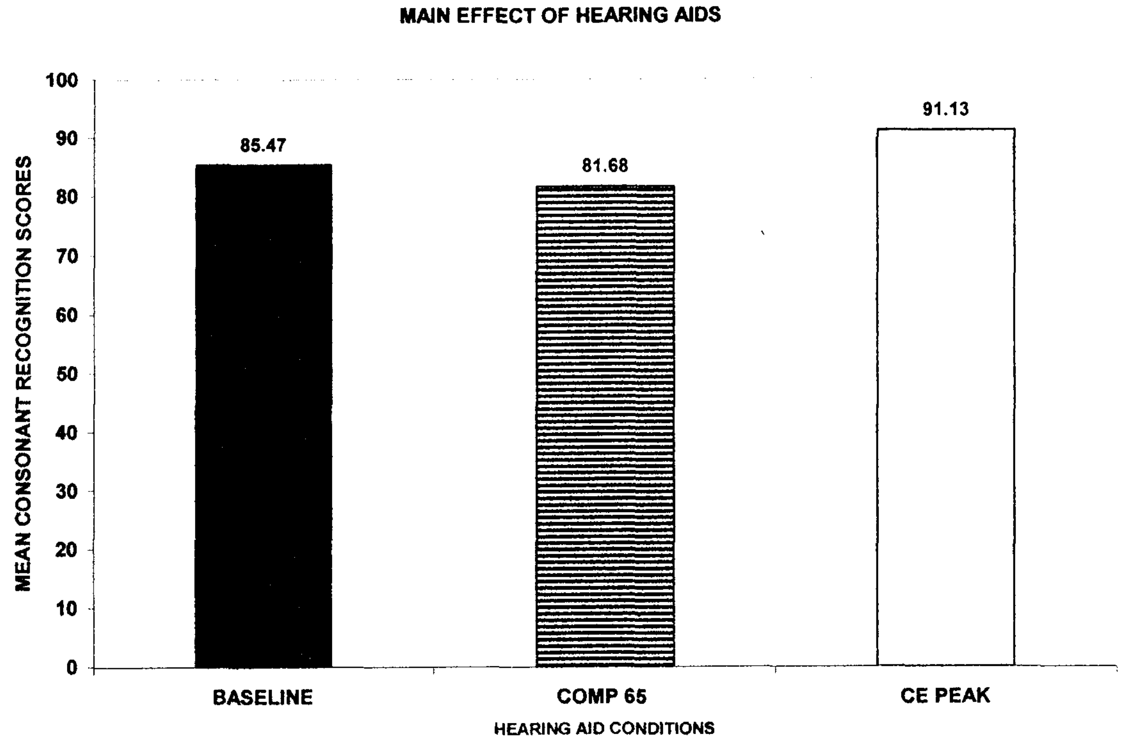


Figure 5. Mean consonant recognition scores as a function of Hearing Aid condition.

The main effect of consonant is illustrated in Figure 6. Tukey post hoc analyses ($p < 0.01$) revealed that mean scores were divided into two groups. Mean scores for /s/, /sh/, /t/ and /k/ did not differ significantly from each other. Mean scores for /f/, /p/ and /th/ did not differ significantly from each other. The scores of the first group were significantly higher than scores for the second group. The mean scores were highest for /s/ (98%) and poorest for /th/ (44%).

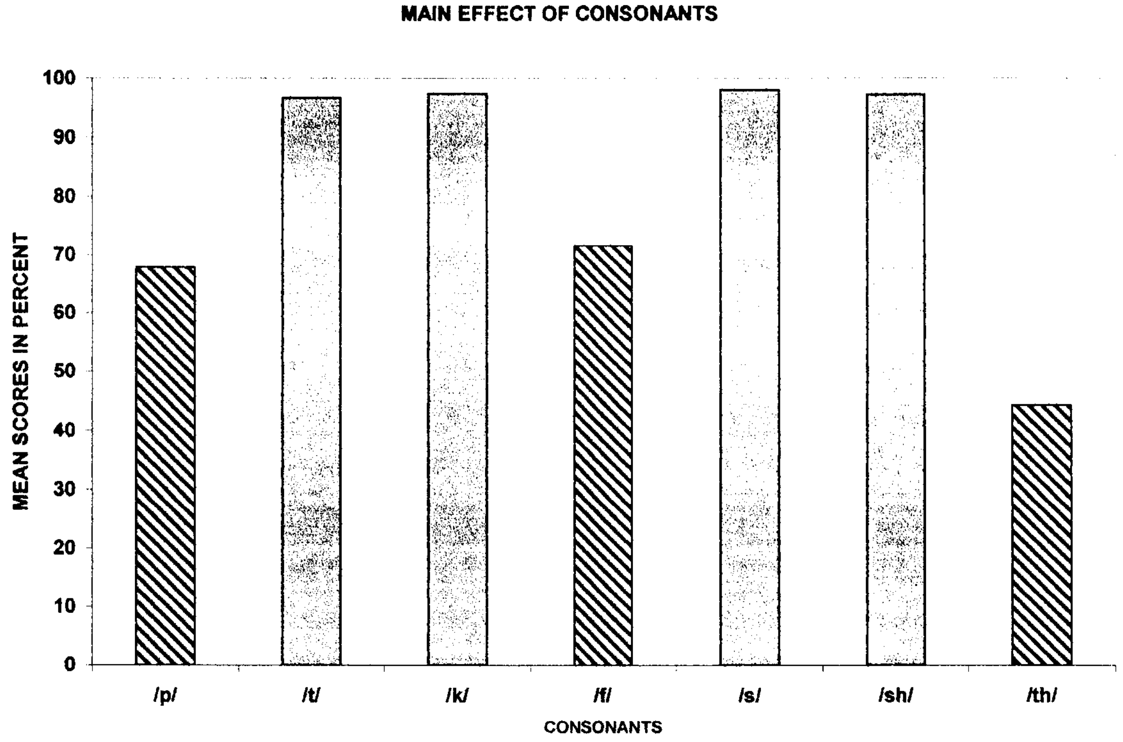


Figure 6. Mean consonant recognition scores of all phonemes in the comfortable listening conditions.

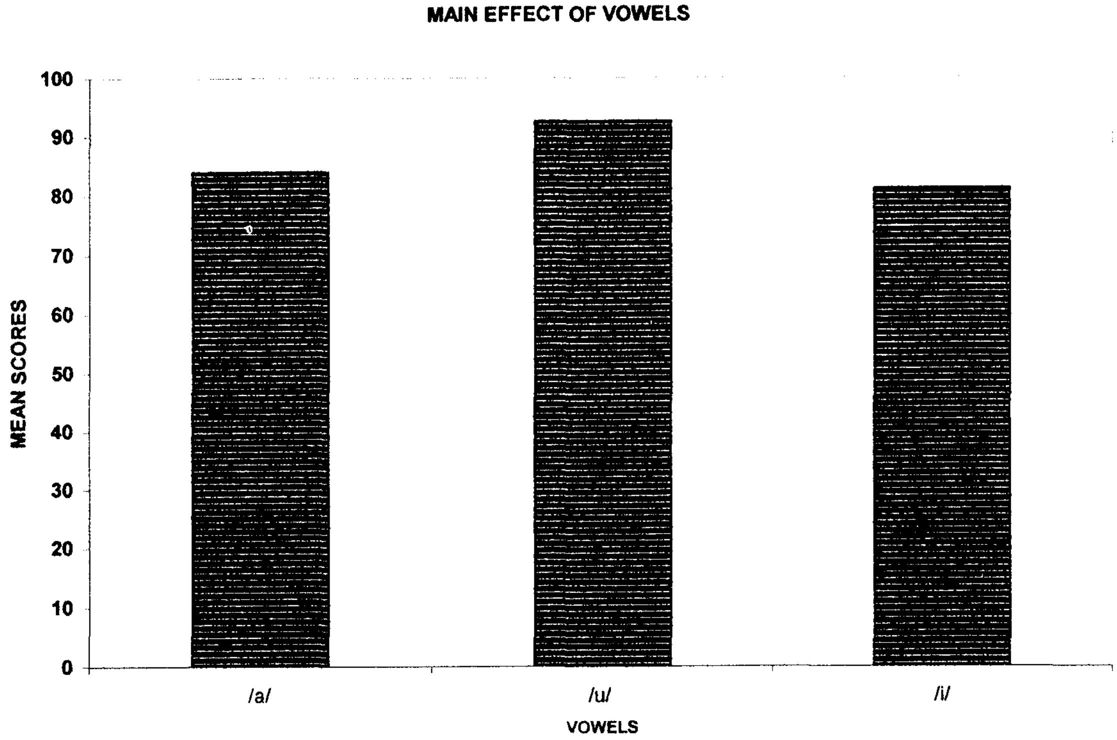


Figure 7. Mean consonant recognition scores as a function of vowel environment.

The mean consonant recognition scores as a function of vowel environment appear in Figure 7. The Tukey post hoc analysis revealed that consonant recognition was significantly better in the /u/ environment (92%), than in the /a/ (84%) and /i/ (81%) environments.

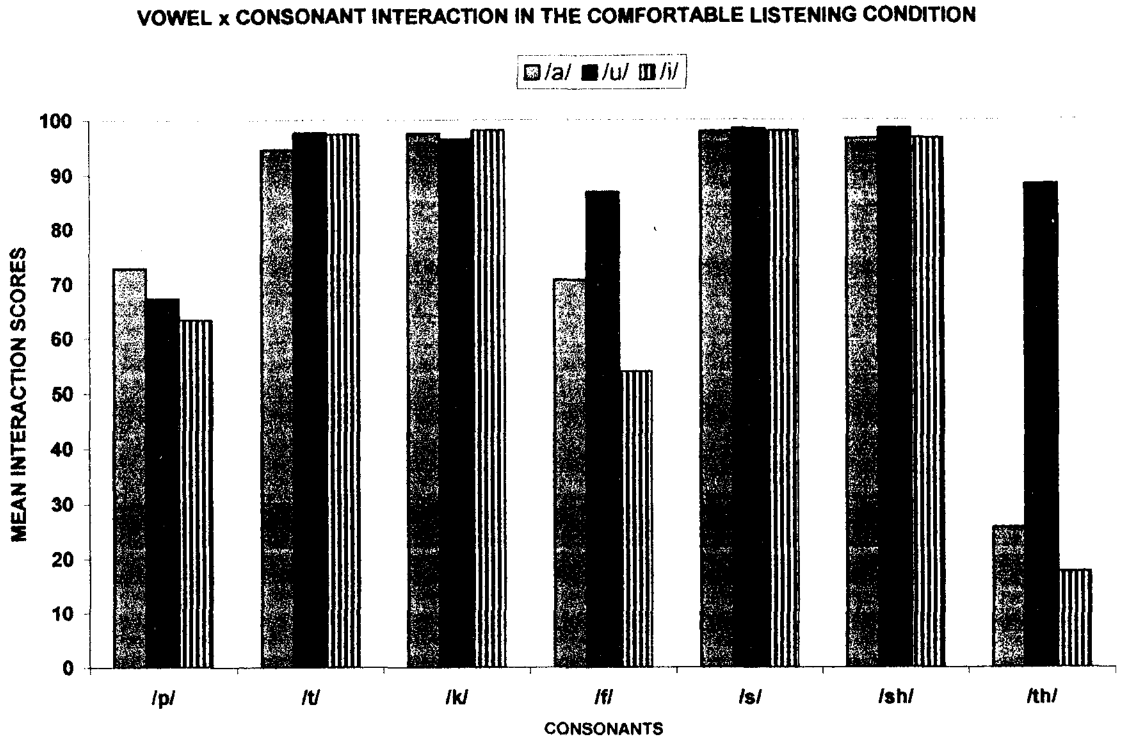


Figure 8. Vowel x Consonant Interaction in the Comfortable Listening Condition

The interaction between consonant and vowel environment is shown in Figure 8. Scores for Consonants /t/, /k/, /s/ and /sh/ were excellent in the context of all three vowels with scores reaching 97%. In contrast, the consonants /f/ and /th/ show differences in performance as a function of the vowel environment. For both of these consonants performance was best in the /u/ vowel context and was poorest in the /i/ vowel context. Mean score for /p/ however was higher in the context of /a/ (73%) and poorer in the context of /u/ (67%) and /i/ (63%).

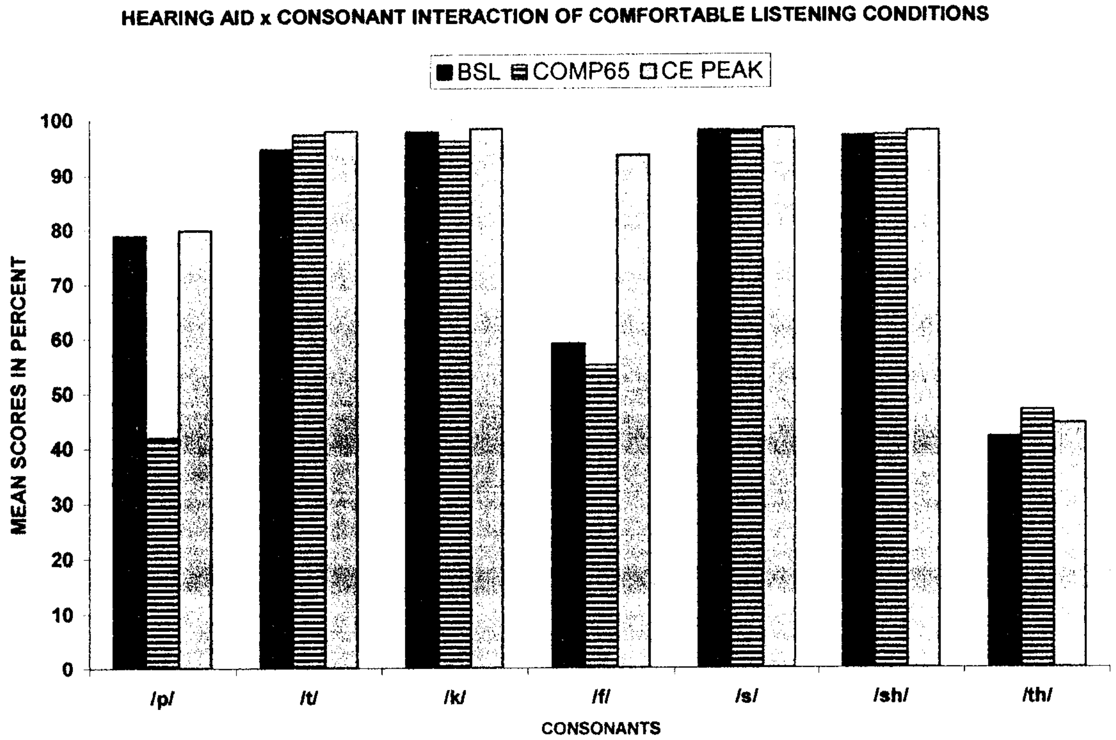
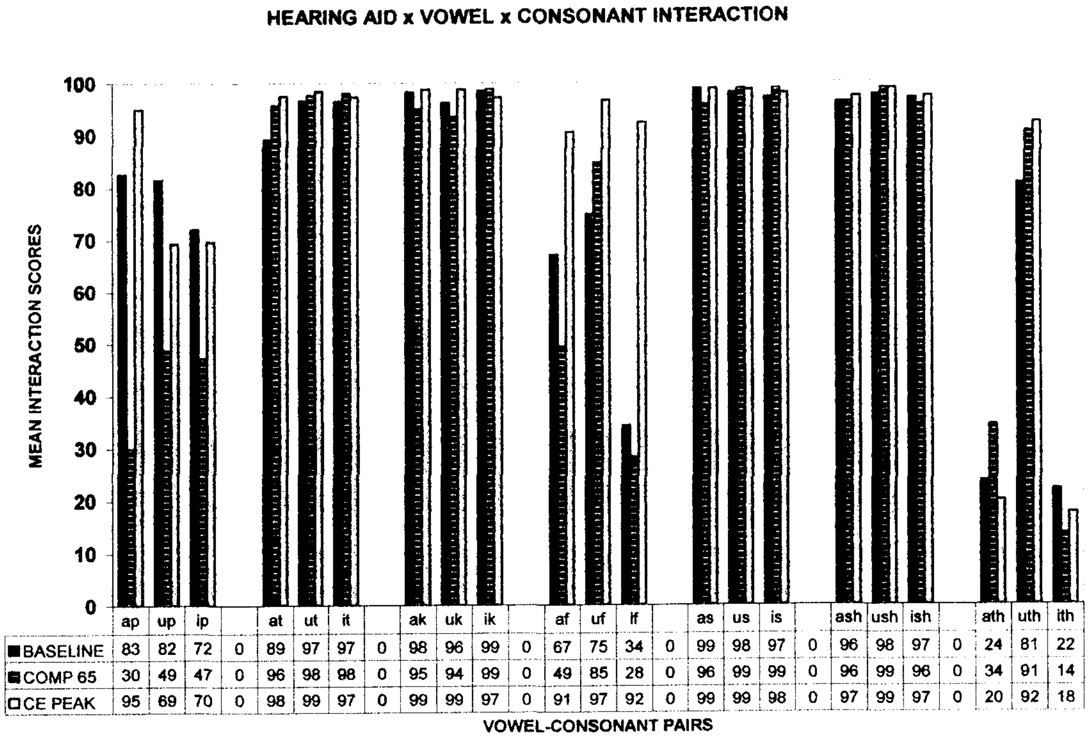


Figure 9.The Hearing Aid x Consonant interaction in the Comfortable Listening Condition.

The significant Hearing Aid x Consonant interaction is illustrated in Figure 9. This interaction shows how the different hearing aid conditions affect consonant perception, which is an integral part of this study. As seen in the figure, the pattern of performance differed as a function of the consonant being amplified. All three hearing aid conditions yielded uniformly high recognition scores for /s/, /sh/, /t/ and /k/. Mean recognition of /p/ was similar in the Baseline and CE Peak condition, but was lower in the COMP65 condition. Mean recognition of /f/ was similar in the Baseline and Compression condition, and was higher in the CE Peak condition. Recognition of /th/ was slightly better in the

Compression condition than the CE Peak condition. The baseline score was the poorest.

Consonant recognition varied significantly ($p < 0.01$) as a function of vowel environment across all hearing aid conditions. Figure 10 shows the three-way interaction of Hearing Aid X Vowel X Consonant. As can be seen in the figure, consonant recognition varies significantly with Hearing Aid conditions and vowel environment. The scores for /t/, /k/, /s/ and /sh/ were uniformly high and showed little variation with Hearing Aid condition or Vowel Environment. In contrast, the scores for /p/, /f/ and /th/ varied in a complex way with Hearing Aid condition and Vowel Environment.



	ap	up	ip		at	ut	it		ak	uk	ik		af	uf	if		as	us	is		ash	ush	ish		ath	uth	ith
■ BASELINE	83	82	72	0	89	97	97	0	98	96	99	0	67	75	34	0	99	98	97	0	96	98	97	0	24	81	22
▤ COMP 65	30	49	47	0	96	98	98	0	95	94	99	0	49	85	28	0	96	99	99	0	96	99	96	0	34	91	14
□ CE PEAK	95	69	70	0	98	99	97	0	99	99	97	0	91	97	92	0	99	99	98	0	97	99	97	0	20	92	18

Figure 10. The Hearing aid x Vowel x Consonant interaction.

Effect of Compression Amplification on C-V ratio and Speech Recognition at Three Different Input Levels.

Compression amplification alters the CV intensity ratio in a complex way that is dependent on the level of the input signal. In the following section, the acoustic and perceptual effects of compression amplification are examined for low, average, and high input levels.

Table 6 identifies the different levels of presentation for each listener under each condition. The input levels were 50, 65 and 80 dB SPL. The compression conditions corresponding to these three input levels are identified as COMP50, COMP65 and COMP80, respectively. The listeners adjusted the gain to MCL for the 65 dB SPL input (COMP65). The gain was then fixed for the COMP50 and COMP80 conditions. The presentation levels for each listener varied for each condition as shown in Table 6. Note that the COMP50 condition is always 16 dB below the COMP65 condition and the COMP80 is 11 dB above COMP65.

Table 6. Presentation levels for each listener (dB SPL) for the three Compression conditions.

	COMP50 dB SPL	COMP65 dB SPL	COMP80 dB SPL
Subject: 1	100	116	127
Subject: 2	90	106	117
Subject: 3	90	106	117
Subject:4	99	115	126
Subject:5	96	112	123
Subject:6	94	110	121
Subject:7	101	117	128
Subject:8	100	116	127
Subject:9	96	112	123
Subject:10	97	113	124
Subject:11	89	105	126
Subject:12	99	115	106

The measured C-V ratios for each of the Vowel-Consonant syllables under the three compression amplification conditions appear in Table 7. The highest C-V ratio for each syllable as a function of Hearing Aid condition is identified with an asterisk.

It was anticipated that the compression conditions would yield higher C-V ratios than the Baseline condition. This observation was found to be valid for the COMP80 condition for all Vowel-Consonant pairs with the exception of /is/, where the Baseline C-V ratio was better by 2.1 dB than the COMP80 condition. Similarly, for the COMP65 condition, the C-V ratios were higher than the Baseline C-V ratios for all but two Vowel-Consonant pairs, /is/ and /ith/. In the Comp50 condition, two-thirds of the syllables had higher C-V ratios than the Baseline condition. Baseline C-V ratios were higher than for the COMP50 condition for seven out of the 21 test syllables; all syllables involved consonants in the context of /u/. The magnitude of C-V ratio change was the largest for /ith/ and smallest for /ak/. Note that the Baseline C-V ratio is higher for/th/ and /f/ in the context of vowels /a/ and /i/.

Table 7. Comparison of Baseline C-V ratio and C-V ratio for Compression at three input levels.

Syllable	BSL	C50	C65	C80
ap	-28	-16.24	-14.1	-11.16 *
ip	-22.37	-26.58	-21.3	-15.64 *
up	-24.86	-7.48	-5.46	-4.78 *
at	-24.16	-16.14	-13.39	-11.33 *
it	-14.22	-9.59	-8.69	-7.62 *
ut	-20.93	-3.63	-2.45	-2.28 *
ak	-25.73	-26.46	-19.64	-15.88 *
ik	-15.47	-12.26	-11.15	-8.81 *
uk	-24.58	-11.31	-9.08	-8.01 *
af	-28.42	-29.28	-22.45	-19.74 *
if	-21.57	-24.79	-21.28	-17.44 *
uf	-21.03	-4.37	-4.21	-3.82 *
as	-11.49	-6.07	-3.86	-2.89 *
is	-2.06*	-5.34	-4.81	-4.14
us	-7.68	13.44*	10.49	7.94
ash	-10.65	-1.94	-1.02	-0.61 *
ish	-10.15	-8.05	-6.95	-4.42 *
ush	-12.11	11.89*	9.05	6.99
ath	-23.52	-27.77	-21.43	-17.29 *
ith	-19.25	-25.61	-20.95	-16.96 *
uth	-23.44	-9.75	-7.99	-7.26 *

Overall, among the compression conditions, COMP80 had the highest C-V ratios followed by COMP65. The C-V ratio for COMP50 was the lowest for all but two syllables, /us/ and /ush/. These two Vowel-Consonant pairs interestingly, had the highest C-V ratios under the COMP50 condition.

Speech Recognition Performance in the Compression Conditions:

The preceding analysis of C-V ratios has shown that compression amplification modifies the C-V ratio differently at different input levels. Since compression amplification is now among the most common signal processing strategies used in hearing aids, it was of interest to examine how well the consonants were recognized at the three input levels. Speech recognition data in proportions were submitted to arc-sine transformation prior to analysis. A repeated measures analysis of variance was completed for all conditions, with Hearing Aid condition (H), Consonant (C), and Vowel (V) as independent variables, and consonant recognition score as the dependent variable. In this analysis, the Hearing Aid conditions correspond to COMP50, COMP65 and COMP80, respectively. The consonants were /p/, /t/, /k/, /f/, /s/, /sh/, and /th/. The vowels were /a/, /i/, and /u/. Statistical significance was set at $p < 0.01$. The Post-Hoc analysis of all significant main effects and interactions was performed using the Tukey Honestly Significant Difference (HSD) test at the $p < 0.01$ level . After the analysis, arcsine values were converted to percentages for ease of interpretation.

A summary of the results of the ANOVA is shown in Table 8. The main effects of Hearing Aid condition, Consonant, and Vowel Environment were all found to be highly significant. Many of the interactions were also found to be significant.

Table 8. Summary Of Repeated Measures Analysis Of Variance For The Three Hearing Aid Conditions

SOURCE OF VARIATION	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARES	F	SIGNF.
H	24.71013	2	12.35507	53.323	0.001
V	12.03523	2	6.01762	21.46	0.001
HV	14.33169	4	3.58292	18.732	0.001
C	330.70898	6	55.11816	32.75	0.001
HC	19.25167	12	1.60431	4.816	0.001
VC	42.35216	12	3.52935	8.526	0.001
HVC	26.41264	24	1.10053	5.92	0.001
S	34.77456	11	3.16132	17.005	0.001
HS	5.09742	22	0.2317		
VS	6.16913	22	0.28041		
HVS	8.41618	44	0.19128		
CS	111.07875	66	1.68301		
HCS	43.9681	132	0.33309		
VCS	54.63942	132	0.41394		
HVCS	49.07793	264	0.1859		
TOTAL	783.02399	755			

The mean consonant recognition scores as a function of the three Hearing Aid conditions are shown in Figure 11. The highest scores are obtained in the COMP65 and COMP80 conditions. A Tukey post-hoc analysis revealed that the mean score in the COMP50 condition (65%) was significantly poorer than the scores obtained in the COMP65 and COMP80 conditions ($p < 0.01$). The mean scores in the COMP65 and COMP80 condition did not differ significantly from each other.

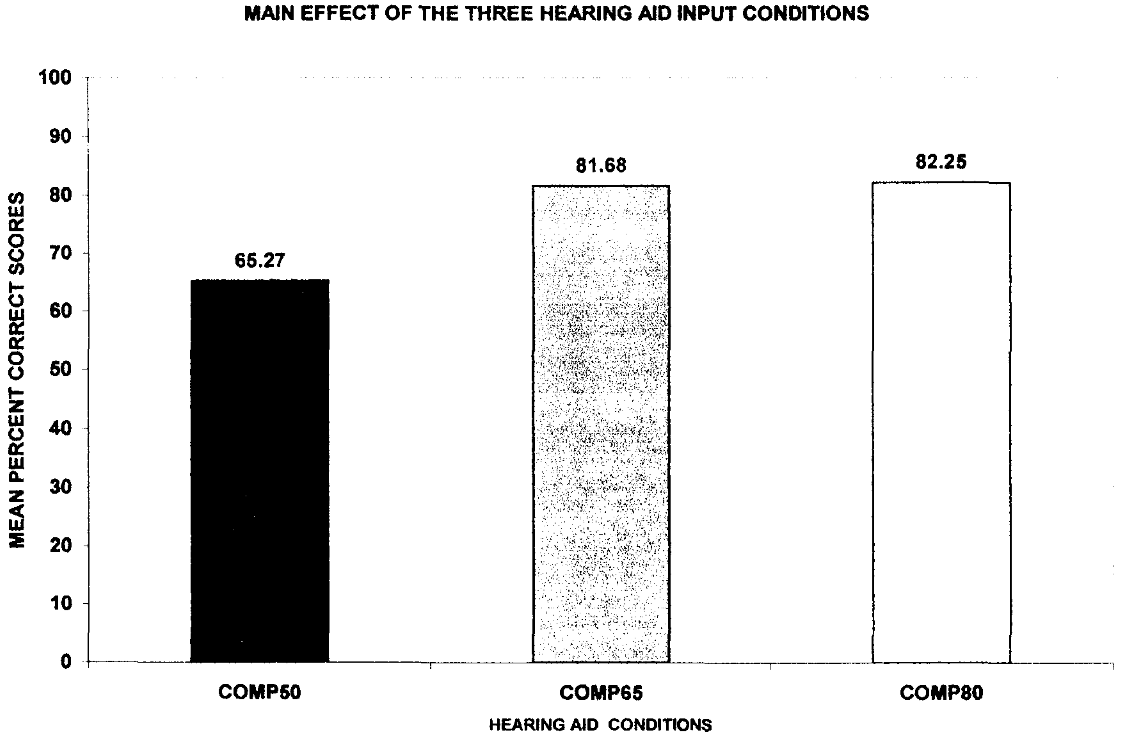


Figure 11. Mean Consonant Recognition Scores as a function of the three Compression Conditions.

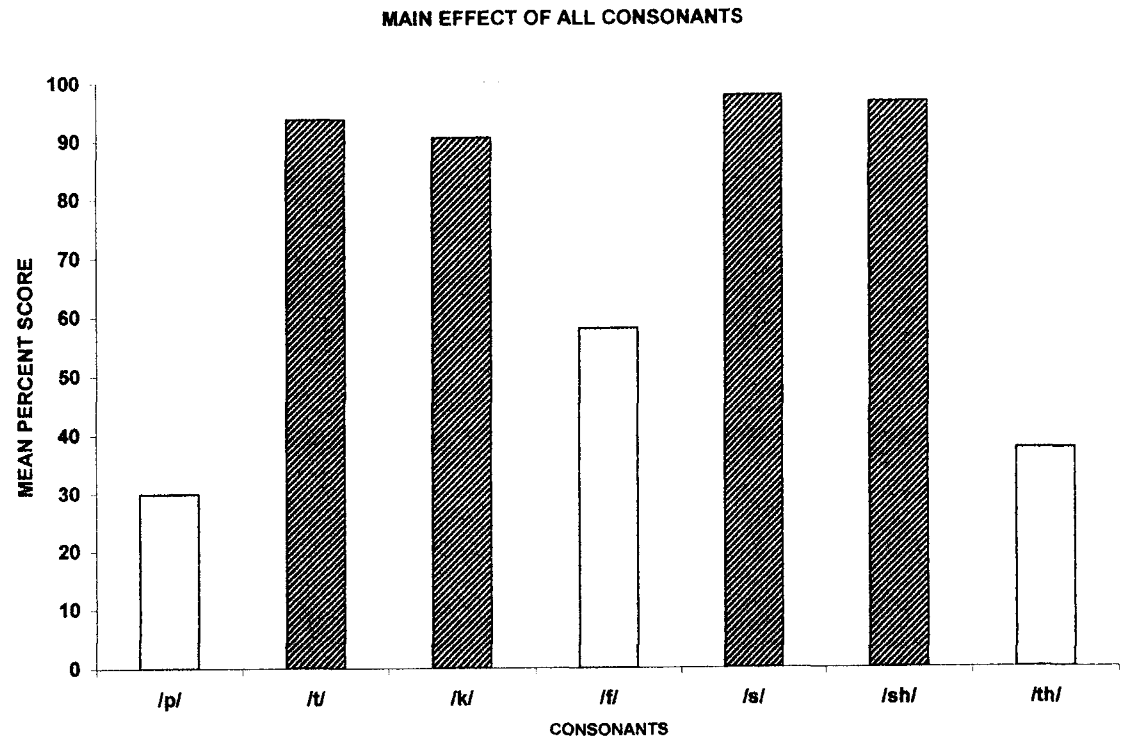


Figure 12. Mean Scores in all Three Hearing Aid (Compression) Conditions as a Function of Consonant.

The main effect of consonants is illustrated in Figure 12. Tukey post hoc analyses ($p < 0.01$) revealed that mean scores were divided into two groups. Mean scores for /s/, /sh/, /t/ and /k/ did not differ significantly from each other. Mean scores for /f/, /p/ and /th/ did not differ significantly from each other. The scores of the first group were significantly higher than scores for the second group. The mean scores were highest for /s/ (98%) and poorest for /p/ (30%).

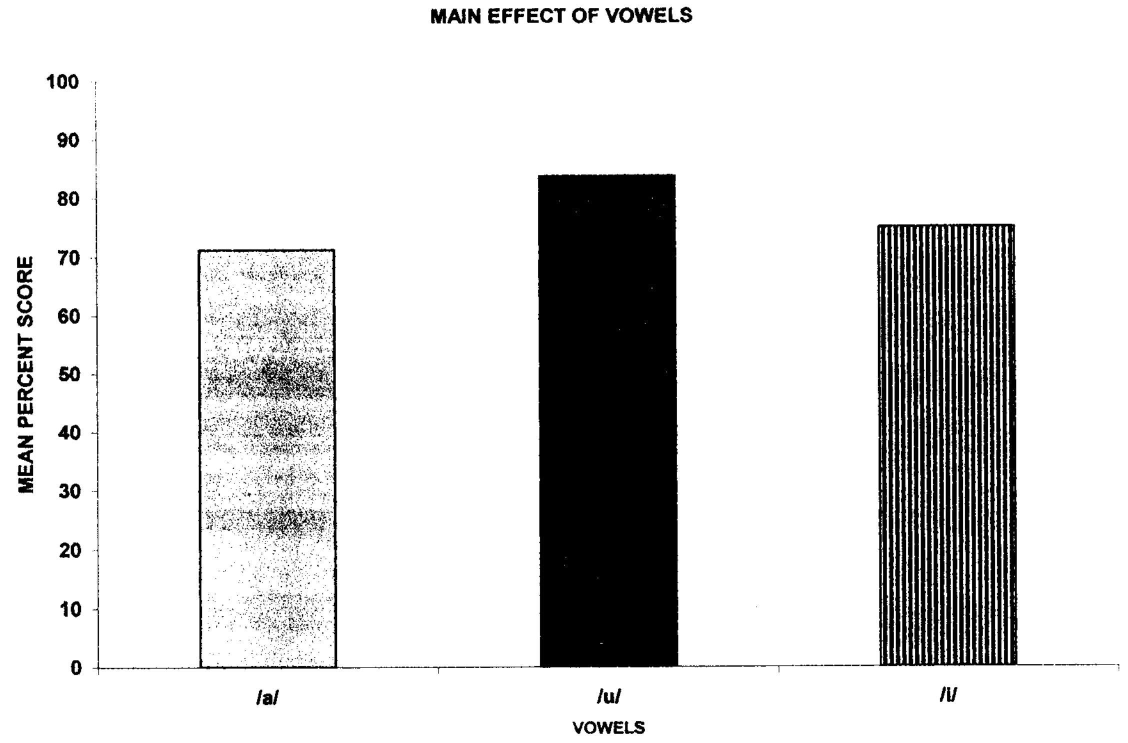


Figure 13. Mean Scores in all Three Hearing Aid (Compression Conditions) as a Function of Vowel Environment.

The mean consonant recognition scores as a function of vowel environment appear in Figure 13. The Tukey post hoc analysis revealed that consonant recognition was significantly better in the /u/ environment (84%), than in the /a/ (71%) and /i/ (75%) environments.

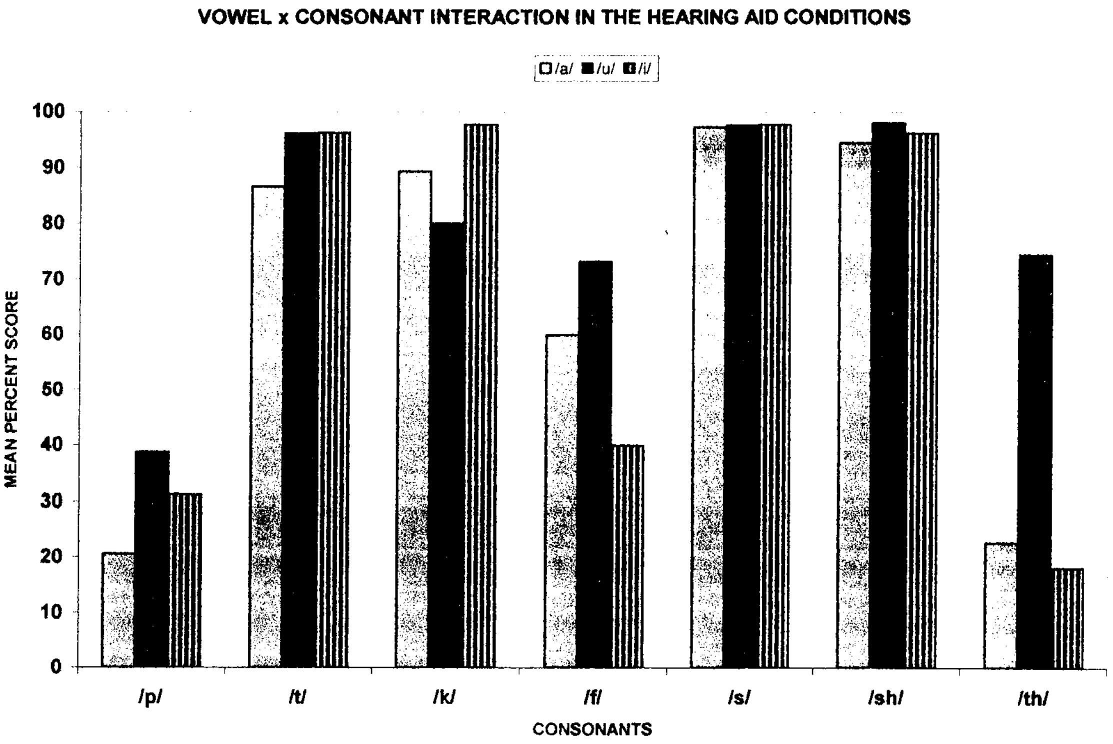


Figure 14. Mean Consonant Recognition for each of the Vowel Environments in all Three Hearing Aid (Compression) conditions.

The interaction between consonant and vowel environment was found to be significant ($p < 0.01$). This interaction is illustrated in Figure 14. The consonants /p/, /f/, and /th/ showed the same pattern of performance, although the absolute level of performance differed among consonants. For these three consonants the best performance was obtained in vowel /u/ context and poorest performance with the vowel /i/. For the consonant /th/, a score of 18% was obtained in the context of /i/, 23% in the context of /a/, and 76% in the context of /u/. The consonant /f/ shows a similar pattern of performance, but different in magnitude. A score of 40% was obtained in the context of /i/, 60% in the context of /a/, and 73% in the context of /u/. Differences in scores as a function of vowel context were smaller for /p/. A score of 21% was obtained in the context of /a/, 31% in the context of / i/, and 39% in the context of /u/.

For the consonant /t/, similar scores were obtained in the context of /i/ and /u/ (96%). Recognition in the context of /a/ was lower (87%).

A different pattern of performance was evident for the consonant /k/; it is the only consonant that had scores decline in the context of /u/. Scores in the context of /i/ and /a/ were 98% and 89%, respectively. The recognition score for /k/ in the context of /u/ was 80%.

Scores for consonants /s/ and /sh/ were highest in the context of all three vowels (95-98%).

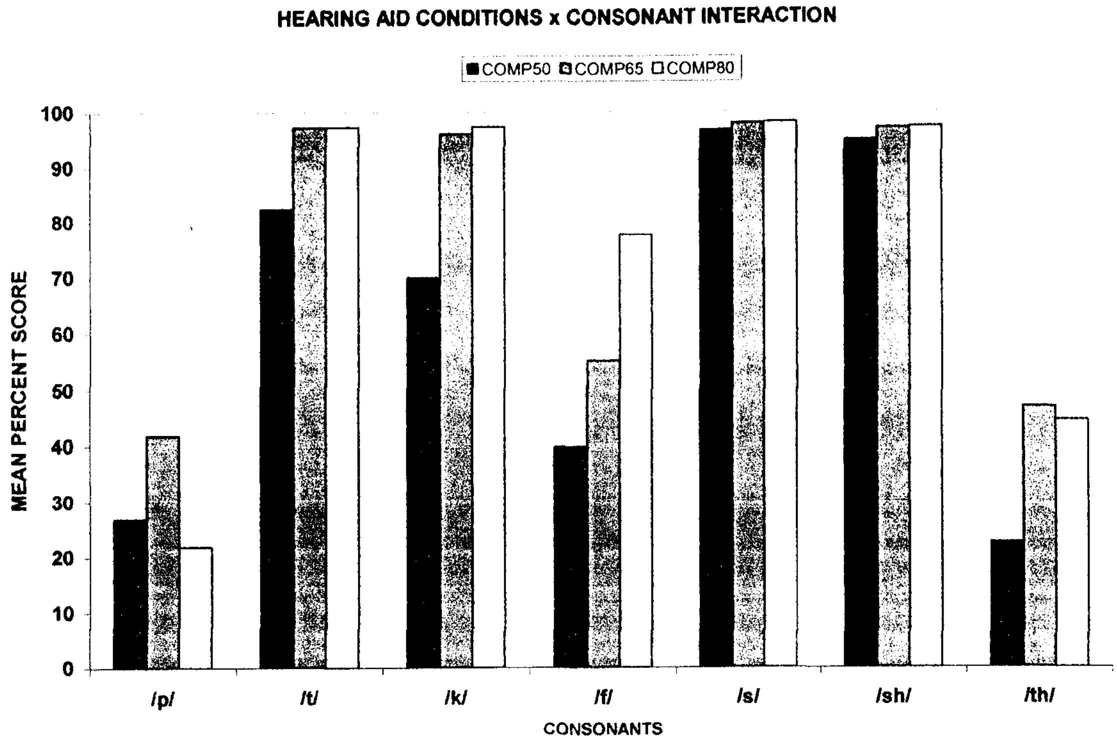


Figure 15. Mean Recognition Score for each Consonant as a Function of Three Hearing Aid (Compression) conditions.

The significant Hearing Aid x Consonant interaction is illustrated in Figure 15. As seen in the figure, the pattern of performance differed as a function of the consonant being amplified. For /s/, and /sh/, all three compression conditions yielded uniformly high recognition scores. For /t/, /k/, and /th/, the lowest score was obtained at the 50 dB input level and similar scores were obtained at 65 and 80 dB input levels. For /f/, mean scores increased with increasing input level. For /p/, the highest score was obtained in the COMP65 condition with lower scores for lower and higher input levels.

The Hearing Aid x Vowel x Consonant interaction is shown for consonants /p/, /t/, /k/ /f/ and /th/ in Figures 16a-16e. Results for /s/ and /sh/ are not shown because scores for these two consonants were uniformly good.

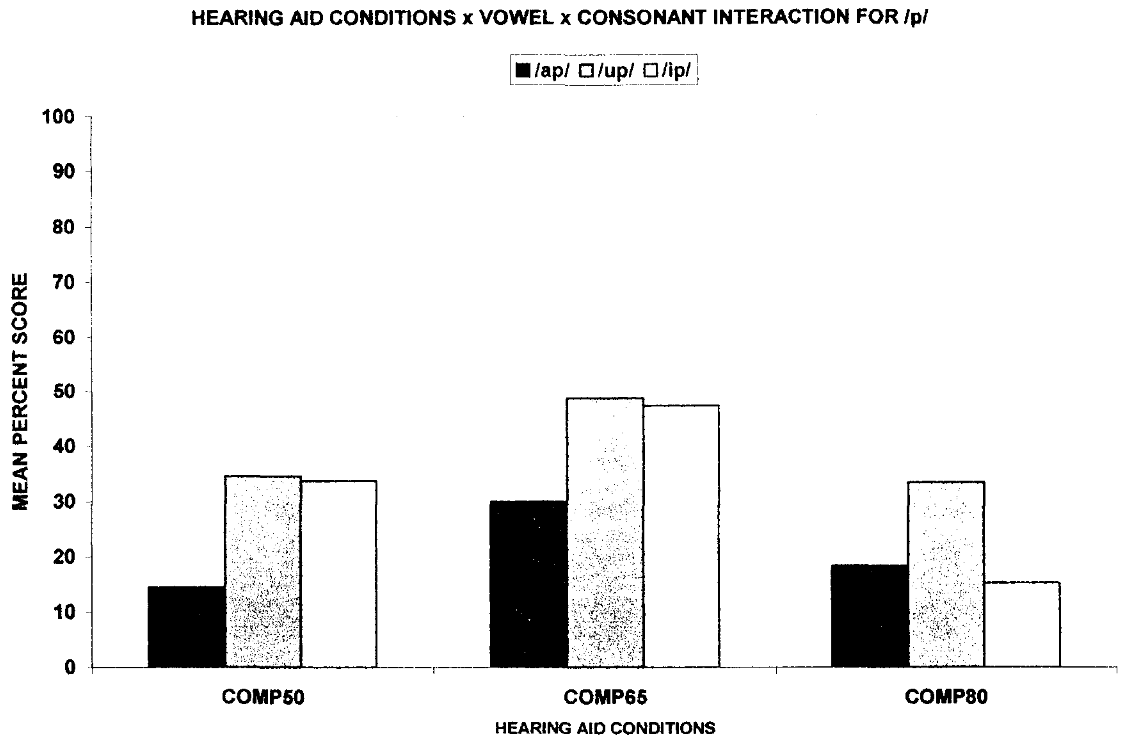


Figure 16a. Recognition of /p/ as a Function of Vowel Environment and Three Hearing Aid (Compression) conditions.

Figure 16a illustrates how recognition of /p/ changes as a function of vowel environment and input level. On average, recognition scores for /p/ were relatively poor. The pattern of performance was similar at the two lower input levels (COMP50 and COMP65), with the poorest performance in the vowel /a/

context. At the highest input level (COMP80), performance was best with vowel /u/ and was similar with /a/ and /i/. The COMP65 condition showed the highest scores in all three vowel environments.

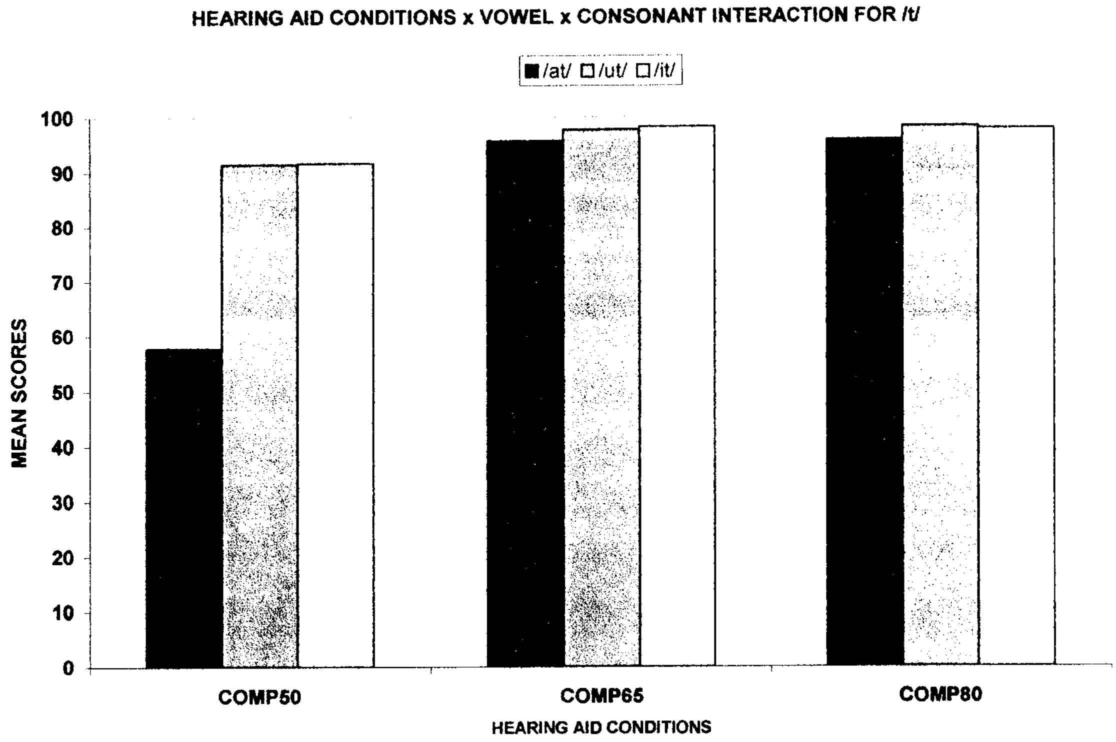


Figure 16b. Recognition of /t/ as a Function of Vowel Environment and Three Hearing Aid (Compression) conditions.

The three way interaction for /t/ is shown in Figure 16b. Mean recognition scores for /t/ were close to 100 % in all three vowel environments at the 65 and 80 dB input levels. Mean scores for the COMP50 condition were almost as good for the /u/ and /i/ vowel contexts (91%), but relatively poor in the /a/ vowel context (58%).

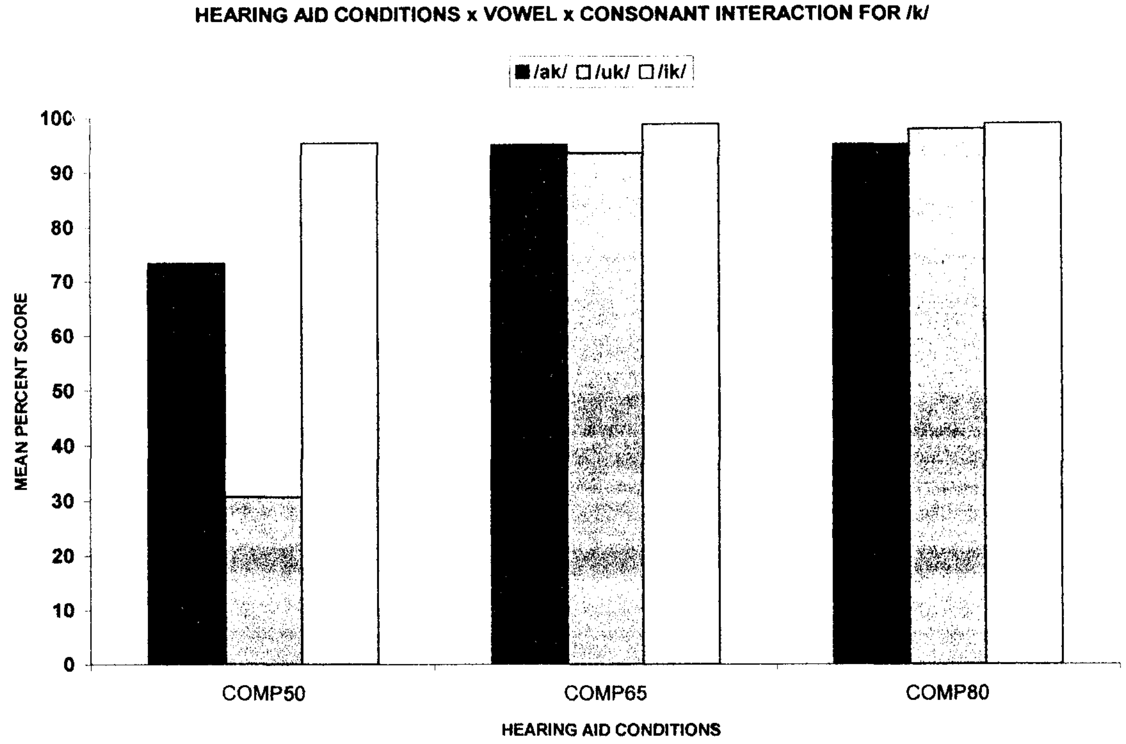


Figure 16c. Recognition of /k/ as a Function of Vowel Environment and Three Hearing Aid (Compression) conditions.

The three way interaction for /k/ is shown in Figure 16c. Similar patterns of performance are noted for the COMP65 and COMP80 conditions where performance was in the upper ninety percent range for all three vowel environments. Scores for the COMP50 condition were much lower for the /a/ and /u/ vowel contexts but relatively good (95%) for the /i/ vowel context.

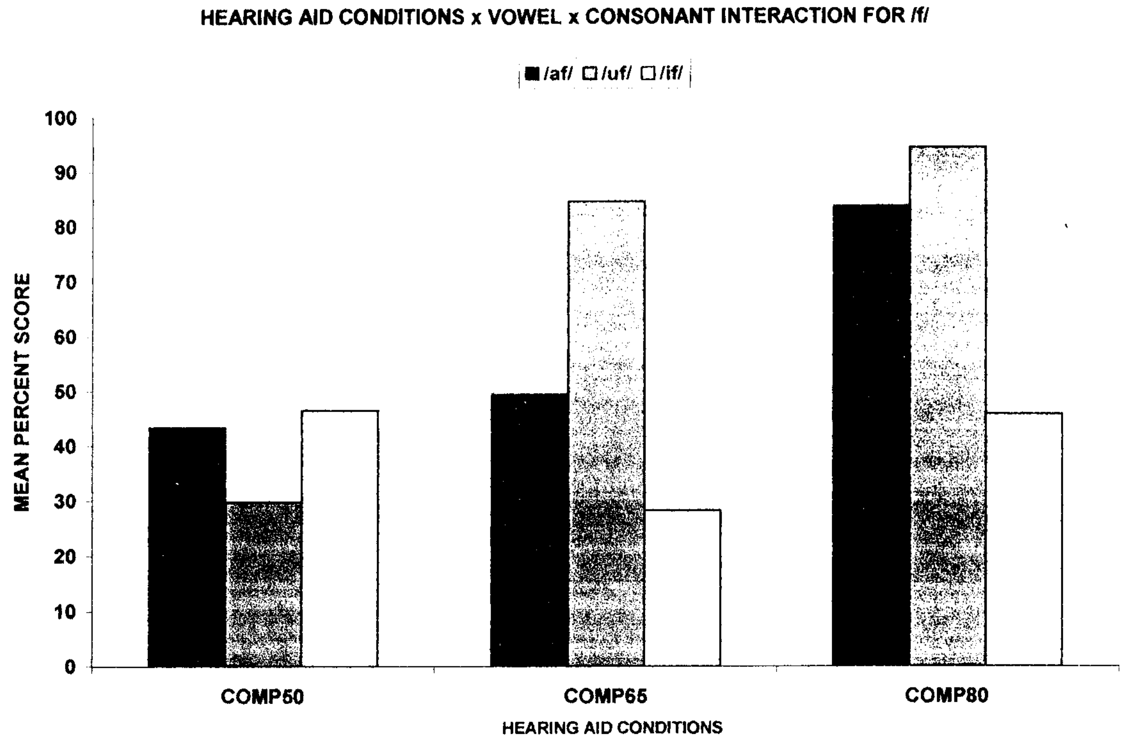


Figure 16d. Recognition of /f/ as a Function of Vowel Environment and Three Hearing Aid (Compression) conditions.

The three way interaction for /f/ is shown in Figure 16d. Again, the pattern of performance is similar at the two highest levels. The best performance is obtained in the /u/ environment and poorest performance in the /i/ environment at the two higher input levels. In contrast, at the 50 dB input level recognition scores are very low (on the order of 20 percent) for all three vowel contexts.

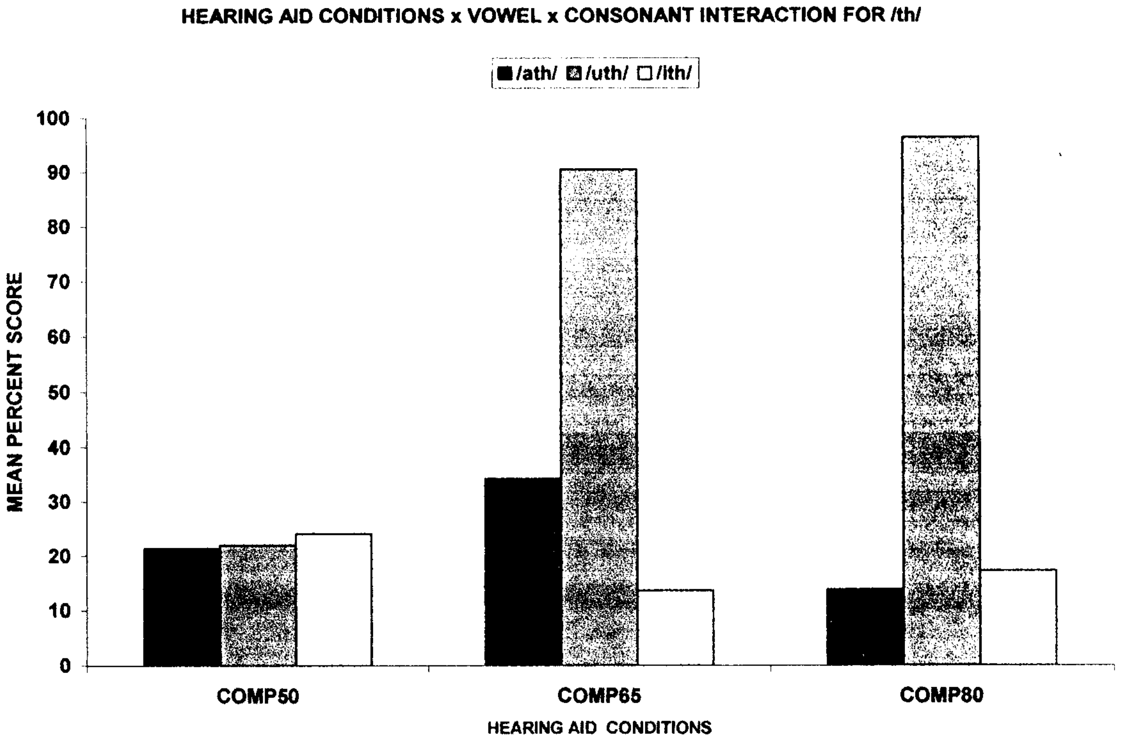


Figure 16e. Recognition of /th/ as a Function of Vowel Environment and Three Hearing Aid (Compression) conditions.

The three way interaction for /th/ is shown in Figure 16e. The pattern of performance is different at each input level. Scores for consonant /th/ were the lowest, on average, although relatively good scores were obtained for the /u/ vowel context in the COMP65 and COMP80 conditions. For the remaining conditions, the recognition scores were in the range of fifteen to twenty percent.

ACOUSTIC ANALYSIS:

Relative levels of Consonants s and Vowels for the five hearing aid conditions:

The RMS levels of the consonant and of the vowel in each syllable were measured using SoundForge™ software. For purposes of the following discussion, these data were normalized such that "0" (full-scale reading) was set to 100 dB (by adding 100 to all values) so that all consonant and vowel values would be positive numbers.

Note that the actual level of presentation for each listener is based on the measured MCL for the carrier phrase and so actual levels of presentation will vary among listeners. Audibility within the group of listeners will vary depending upon the MCL setting for each listener.

The RMS values (relative to full scale) for each vowel and consonant, corresponding C-V ratios, and consonant recognition scores appear in Appendix 2a-c. The measured vowel and consonant levels reveal that compression at different input levels modifies the consonant and vowel levels differentially (and will presumably also modify audibility differentially, although this was not measured). The consonants /p/, /f/ and /th/ showed substantial changes in their C-V ratio as a result of compression.

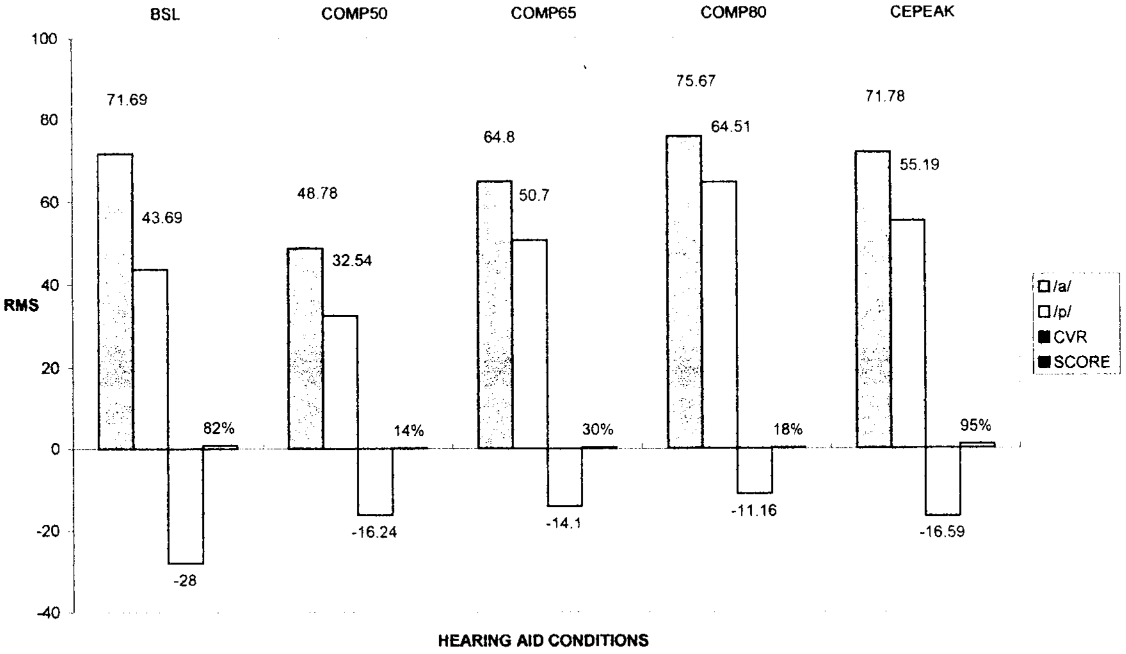


Figure 17: Relationship between Relative Vowel-Consonant values and C-V ratio and Corresponding Consonant Recognition Score for /ap/ in all five Hearing Aid Conditions.

The relative vowel-consonant RMS values for the baseline, the three compression, and the CE Peak conditions for /ap/ are illustrated in figure 17. Tables for all three consonants in the context of the three vowelsin Appendix 2a-2c.

In figure 17, for the Baseline condition, the consonant level was 43.7 dB and the vowel level was 71.7 dB yielding a C-V ratio of -28.0 dB. For the COMP50 condition, the vowel level was reduced by 22.9 dB to 48.8 dB while the consonant level was reduced by only 9.2 dB to 32.5 dB, thereby increasing the C-V ratio to -16.2 dB; the C-V ratio is increased while the overall sound level is decreased. For the COMP65 condition, the input level is increased by 15 dB (relative to COMP50). The vowel level is increased by almost the same amount (16 dB) to a new level of 64.8 dB. The consonant level, however, is increased by a greater amount (18.2 dB) to a new level of 50.7dB. The net effect is to increase the C-V ratio by 2.1 dB to -14.1 dB (18.2 – 16.1 = 2.1).

Compared to the COMP65 condition, the COMP80 condition vowel level was increased by 10.9 dB to 75.7 dB, while the consonant level was increased by 13.8 dB to 64.5 dB. These changes increased the C-V ratio to -11.2 dB. In this case, both the C-V ratio and overall sound levels are increased. The net effect is to increase the C-V ratio by 2.94 dB to -11.16 dB (13.81 – 10.87 = 2.94).

Finally, compared to the COMP80 condition, the CEPEAK condition vowel level was decreased by 3.89 dB to 71.78 dB, and the consonant level was decreased by 9.32 dB to 55.19 dB. These decreases in overall sound levels led to the C-V ratio decreasing 5.34 dB to -16.59 dB. (-9.32-(-3.98) = -5.34).

In order to determine how consonant recognition scores change with C-V ratio, a ratio of the difference in consonant recognition scores (from Baseline condition) to the difference in C-V ratio (from baseline) was calculated for each syllable for each listener for the COMP65 and CE Peak conditions. A repeated measures analysis of variance was carried out on these ratios to determine the effect of Hearing Aid (COMP65 and CE Peak), Vowel (a, u, i), and Consonant (p, f, th). The factor of Hearing Aid was found to be significant (F = 4.97, df =1, 11, p

< 0.05). The mean ratio score for the COMP65 condition was –7.07. The mean ratio for the CE Peak condition was 1.15. These values indicate that for the COMP65 condition the mean relationship between recognition and C-V ratio was negative (i.e., no increase in recognition with increasing C-V ratio), while the relationship between recognition and C-V ratio for the CE Peak condition was positive (i.e., recognition increased with increasing C-V ratio). None of the other main factors or interactions was significant.

The mean and standard deviation of the ratios appear in Table 9. Negative numbers indicate a negative relationship (slope) between recognition and C-V ratio. Positive numbers indicate a positive relationship, i.e., increases in C-V ratio are associated with increases in consonant recognition. Examination of the table reveals that, for the most part, in the COMP65 condition the mean values are negative or 0, while in the CE Peak condition all of the mean values are positive.

Table 9. Means and standard deviations of the ratios of the differences between Baseline and COMP65 and Baseline and CE Peak for /p/, /f/ and /th/ in the context of /a/, /u/ and /i/.

/ap/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	-3.05755		Mean	1.351154
	Std Dev	3.073383		Std Dev	3.078609
/up/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	-1.33162		Mean	-0.77986
	Std Dev	1.709012		Std Dev	2.142453
/ip/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	-18.3022		Mean	-0.25376
	Std Dev	21.9593		Std Dev	4.48629
/af/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	-2.1636		Mean	2.285671
	Std Dev	6.107259		Std Dev	3.267577
/uf/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	0.47067		Mean	1.980568
	Std Dev	1.683774		Std Dev	2.444654
/if/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	-47.4138		Mean	3.165915
	Std Dev	108.0795		Std Dev	2.534476
/ath/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	0.99681		Mean	1.02459
	Std Dev	12.47103		Std Dev	3.998833
/uth/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	0.977836		Mean	1.530456
	Std Dev	1.632687		Std Dev	2.784803
/ith/	65 - Baseline			CEpeak - Baseline	
	Difference	slope		Difference	slope
	Mean	6.127451		Mean	0.039683
	Std Dev	17.17086		Std Dev	4.198037

Analysis of consonant confusion:

Confusion matrices for all of the experimental conditions are shown in Appendices 3a-3e.

The following tables (10-12) summarize the response patterns in all five hearing aid conditions for each VC. Since listeners' responses were based on one out of seven choices, the chance of guessing correctly is placed at 14.3%. Hence, only the responses over 15% have been tabulated. Because errors for /s/ and /sh/ were always very low, data for these consonants are not shown in these tables, but only in the appendices.

Table 10: Responses (in percent) in the context of /a/ for the five hearing aid conditions.

	Target	Response						
Condition								
	af	af	ak	ap	as	ash	at	ath
BSL		60		15				24.2
COMP50		37.9		29.6				17.9
COMP65		47.1						40.8
COMP80		76.7						
CE Peak		73.75						24.17
	ak	af	ak	ap	as	ash	at	ath
BSL			99.2					
COMP50			70					
COMP65			91.7					
COMP80			91.7					
CE Peak			99.5					
	ap	af	ak	ap	as	ash	at	ath
BSL				76.25			15.42	
COMP50		15.4		18.3			48.3	15.8
COMP65				35.4			61.3	
COMP80				23.3			70.8	
CE Peak				75			25	
	at	af	ak	ap	as	ash	at	ath
BSL		16					84.16	
COMP50							57.1	18.3
COMP65							93.8	
COMP80							94.2	
CE Peak							96.25	
	ath	af	ak	ap	as	ash	at	ath
BSL		69.6						24.58
COMP50		33.3		22.9			17.5	17.5
COMP65		64.6						32.9
COMP80		81.3						12.9
CE Peak		83.75						15

Inspection of Table 10 reveals that in the vowel /a/ context, most errors were made in recognition of the consonants /f/, /p/, and /th/. The most common substitution for /f/ was /th/; the most common substitution for /p/ was /t/, and the most common substitution for /th/ was /f/.

The errors for the three compression hearing aid conditions differed from the Baseline and CE Peak conditions. In the context of /a/, substitution of /p/ for /f/ was 30% in the COMP50 condition, while in the COMP65 condition it was only 8% and in the COMP80 condition, this substitution was not evident. On the other hand, the /t/ for /p/ substitution was evident across all three compression conditions, with percentage of errors progressing from 48% in the COMP50 condition, to 61% in the COMP65 condition, to 71% in the COMP80 condition. The /th/ for /f/ substitution was evident mainly, in the COMP65 condition; substitution of /th/ for /f/ was 40%, and /f/ for /th/ was 65%. Substitution of /t/ for /p/ was 61%.

Confusions in the Baseline and CE Peak conditions were predominantly in the context of /f/ and /th/, where /ath/ for /f/ was 24% under both these hearing aid conditions but /a/ for /ath/ was 69% in the Baseline condition and 84% in the CE Peak condition.

The following table (Table 11) shows the confusion matrices in all five hearing aid conditions for each VC syllable in the context of /u/.

Table 11: Responses (in percent) in the context of /u/ for the five hearing aid conditions.

	Target	Response						
Condition								
	uf	uf	uk	up	us	ush	ut	uth
BSL		71.25						15.42
COMP50		22.9					35.4	20.4
COMP65		80.4						19.6
COMP80		90						
CE Peak		93.3						
	uk	uf	uk	up	us	ush	ut	uth
BSL			93.3					
COMP50		23.8	34.6	23.3				
COMP65			90					
COMP80			97.9					
CE Peak			99.6					
	up	uf	uk	up	us	ush	ut	uth
BSL				74.58			19.17	
COMP50		16.3		38.8			20	
COMP65				48.3			39.6	
COMP80				37.5			51.3	
CE Peak				63.75			28.75	
	ut	uf	uk	up	us	ush	ut	uth
BSL							95	
COMP50							87.5	
COMP65							97.5	
COMP80							98.8	
CE Peak							99.17	
	uth	uf	uk	up	us	ush	ut	uth
BSL								72.92
COMP50		37.5						28.8
COMP65								87.9
COMP80								97.9
CE Peak		16.25						83.75

As shown in Table 11, the majority of errors in the /u/ context were for the consonants /f/, /p/, and /th/. The most common substitution for /f/ was /th/, the most common substitution for /th/ was /f/, and the most common substitution for /p/ was /t/. The /p/ for /t/ confusion occurred only in the compression conditions. The following table, (Table 12) shows the confusion matrix in all five hearing aid conditions for each VC syllable in the context of /i/.

Table 12: Responses (in percent) in the context of /i/ for the five hearing aid conditions.

	Target	Response				Response		
Condition								
	if	if	ik	ip	is	ish	it	ith
BSL		30.42						66.67
COMP50		40.0						32.5
COMP65		24.2						70.0
COMP80		37.1						56.3
CE Peak		73.75						26.25
	ik	if	ik	ip	is	ish	it	ith
BSL			99.58					
COMP50			100					
COMP65			100					
COMP80			100					
CE Peak			94.58					
	ip	if	ik	ip	is	ish	it	ith
BSL				61.25			30.42	
COMP50		21.7		35.8			19.2	20.4
COMP65				47.5			45.4	
COMP80				20.8			76.3	
CE Peak				64.58			31.67	
	it	if	ik	ip	is	ish	it	ith
BSL							95.0	
COMP50							86.70	
COMP65							98.80	
COMP80							97.50	
CE Peak							94.58	
	ith	if	ik	ip	is	ish	it	ith
BSL		62.50						28.33
COMP50		27.9		20.8				27.9
COMP65		82.5						17.5
COMP80		65.4						26.7
CE Peak		71.67						23.33

As shown in Table 12, error patterns in the /i/ context were similar to those of /a/ and /u/. Substitutions included /th/ for /f/, /f/ for /th/, and /t/ for /p/. Again, the /t/ for /p/ substitution was confined to the compression conditions.

In summary, the most common error pattern was /th/ for /f/ substitution in the Baseline, COMP65, COMP80 and CE Peak conditions. The error pattern in the COMP50 condition was more diverse. Substitution errors were spread across consonants. Target sound /f/ was substituted with /p/ and /th/. Interestingly, consonant substitution of /t/ for /p/ was observed only in the compression conditions. This substitution was larger for COMP80 followed by COMP65. COMP80 had the largest percentage of /f/ for /th/ substitution errors followed by COMP50 errors and CE Peak. The largest percentage of /t/ for /p/ substitution was under COMP80 followed by COMP65, CE Peak and the Baseline. COMP50 errors were again spread across /t/ and /f/. In the context of /i/, COMP50 had the least percentage of errors.

Spectral Analyses:

The error patterns show a systematic /f/ for /th/, /th/ for /f/, and /t/ for /p/ substitution and these errors occurred more in the three compression conditions. Substitution of /f/ for /th/ and /th/ for /f/ can possibly be explained by the fact that when gain is increased for soft sounds, the fricative noise gets amplified, resulting in /f/ for /th/ confusion. In order to obtain a better understanding of why the labial /p/ was perceived as a /t/ after compression, spectrographic analyses were carried out on the consonants /p/ and /t/ in all three vowel contexts for the baseline, COMP65 and COMP80 and CE Peak conditions.

Wideband spectrograms were obtained using the Multi-Speech software (Version 2.7, Model 3700, Kay Elemetrics). Since the largest number of /t/ for /p/ substitution errors occurred in the COMP80 conditions, the spectral analyses of the unenhanced (Baseline) stimuli and the compressed stimuli (COMP80) for the syllables /ap/ and /at/ are illustrated and described here. For comparison, the CE Peak condition is also included. Spectrograms for /p/ and /t/ in the /i/ and /u/ environments are shown in Appendix 6.

The following figure (Figure 18) shows spectrograms of syllables /ap/ and /at/, respectively, in the Baseline, COMP80 and CE Peak conditions. The right side panels show the spectra of /ap/ and the left panels show the spectra of /at/. The top panels show the Baseline condition, the middle panels show the COMP80 condition and the bottom panels show CE Peak condition.

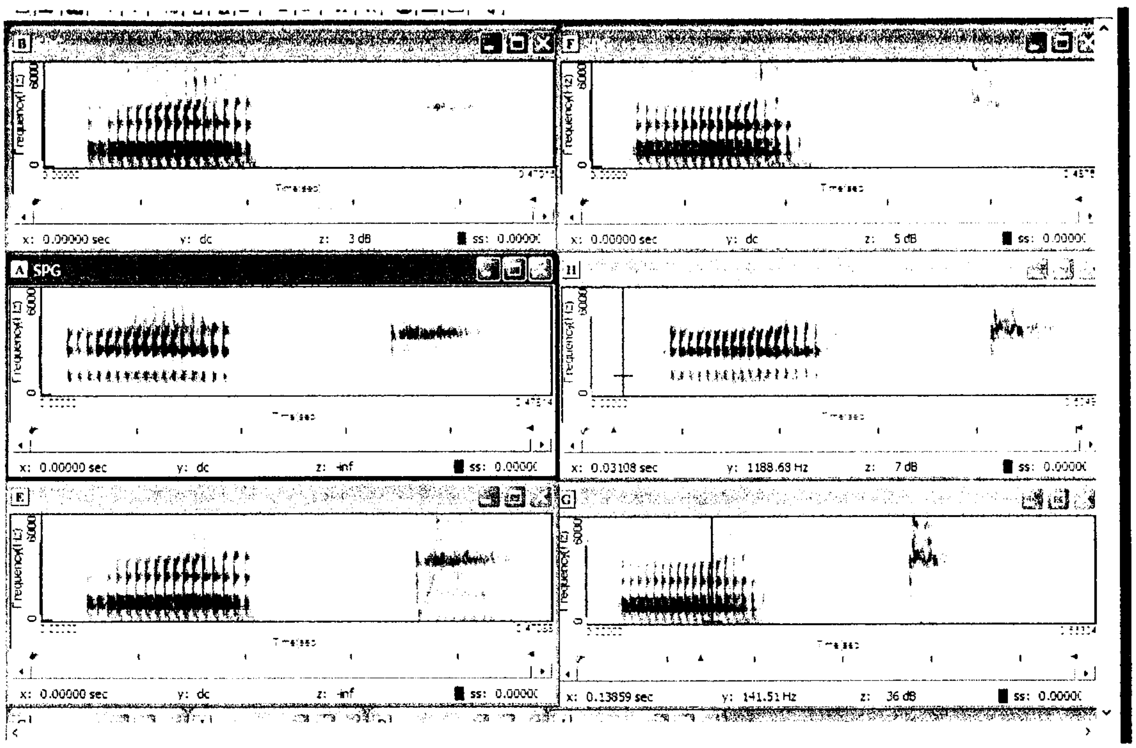


Figure 18. Spectrograms of /ap/ (left panels) and /at/ (right panels) in the Baseline (top), COMP80 (middle) and CE Peak (bottom) conditions.

A comparison of the spectrograms of /ap/ and /at/ in the Baseline conditions reveals more energy in the burst of the /t/ than in the burst of the /p/. The energy in the burst of the /p/ is less than that in the vowel formant in the preceding vowel, whereas the energy in the burst of the /t/ is more similar in level to that of the formant at a similar frequency in the preceding vowel. A comparison of the spectrograms of the syllables /ap/ and /at/ in the COMP80 condition (middle panel) reveals that the bursts in the two syllables look much more alike than in the baseline condition. Both have greater energy than in the baseline condition and have a concentration of energy at a similar frequency. The energy in the burst is similar in level to that in the formant in the same frequency region in both the /p/ and the /t/. Examination of the spectrograms for /at/ and /ap/ in the CE Peak condition (bottom panel) reveals that the differences evident in the baseline /ap/ and /at/ bursts are maintained. Although the increase in level in the burst is evident in both the /p/ and the /t/, the burst energy in the /p/ is lower than in the adjacent vowel formant, whereas the energy in the /t/ is still greater than that in the /p/. It appears that compression changes the amplitude cue necessary for the perception of /p/ in a manner that CE Peak processing does not. This same pattern was seen in the spectrograms for syllables /ip/ and /it/ and /up/ and /ut/ (Appendix 6).

CHAPTER V

Discussion

Past research has shown that enhancement of the C-V ratio can yield improved speech recognition performance (Hecker, 1974; Gordon-Salant, 1985; Montgomery and Edge,1988; Freyman and Nerbonne, 1989; Kennedy, Levitt, Neuman and Weiss,1998, Smith and Levitt,1999). Kennedy and colleagues (1998) demonstrated that for individual listeners, different amounts of consonant enhancement were required as a function of consonant and vowel context in order to obtain maximum consonant recognition. Smith and Levitt (1999) found similar results in children. The findings of these studies suggested that the C-V ratio would need to be adjusted in a very complex manner for different consonants in different vowel contexts and for each listener. On the other hand, syllabic compression amplification provides a relatively straightforward method of obtaining C-V ratio enhancement (Hickson et a,l 1995, Hickson, Thayer and Bates, 1999, Preves, Fortune, Woodruff and Newton 1991, Balakrishnan et al, 1996). Hence one of the aims of the current study was to compare consonant recognition performance obtained using syllabic compression amplification to that obtained using direct C-V ratio enhancement (CE Peak enhancement) at most comfortable listening level. Unprocessed stimuli at most comfortable listening level served as the baseline (reference condition). Test stimuli consisted of nonsense syllables (VC) comprised of seven voiceless consonants in combination with each of three vowels (/a/, /u/, and /i/). In addition, acoustic analyses of the compressed, CE Peak enhanced, and baseline stimuli included measures of the levels of the vowel and consonants and of the C-V ratio.

The main findings of the comparisons between Baseline, COMP65 and CE Peak conditions were:

1. Mean voiceless consonant recognition performance was significantly higher in the CE Peak condition than in the Compression65 or Baseline conditions. Compression and baseline recognition scores did not differ significantly.
2. There was a significant interaction between Hearing Aid condition and Consonants. While performance was equivalent for Compression65, CE Peak, and Baseline for consonants /s/ and /sh/, and /t/ and /k/, the CE Peak condition was clearly best for recognition of /f/, and Compression65 clearly degraded recognition of /p/, while it slightly improved recognition of /th/.
3. Compression65 increased the C-V ratio (from the natural C-V ratio) for all but two of the 21 test stimuli. The CE Peak condition had higher C-V ratios than compression for most of the stimuli, however compression (65 dB input) yielded higher C-V ratios than CE Peak for all consonants in the context of /u/.
4. Compression amplification altered C-V ratio in three ways-- sometimes by decreasing vowel intensity, sometimes by increasing consonant level, and sometimes by doing both. This is in contrast to the changes in C-V ratio by the

method of C-V ratio enhancement where the vowel level remains constant and the consonant level is increased.

Speech Recognition Performance At Comfortable Listening Levels

CE Peak And Baseline Results

A number of other studies have explored the use of C-V ratio modification to maximize consonant intelligibility. Most of these studies demonstrated that increasing the amplitude of consonant levels without altering vowel levels can improve intelligibility, although the magnitude of improvements varied, probably due to procedural differences. Because the same stimuli and similar procedures were used in the present study as in the study by Kennedy and colleagues, it is possible to compare the CE Peak performance obtained in the present study to performance of listeners in the earlier study. Kennedy and colleagues found an average 24% improvement (from the baseline condition of unenhanced stimuli presented at MCL) for voiceless consonant recognition as a result of individualizing consonant enhancement. In the present study, a small, but significant mean improvement of 5.7% was obtained. Thus the finding in the present study of superior performance for the CE Peak condition is in agreement with the earlier work of Kennedy, although the magnitude of improvement is much smaller.

Inspection of the data from the two experiments reveals that the listeners in the Kennedy study started out with much poorer recognition in the baseline condition and thus had much more chance for improvement. In the present study, recognition of the intense fricatives (/s/ and /sh/) started out close to ceiling (98% and 97%, respectively). Initial recognition of the stops /t/ and /k/ were also quite high (95% and 98%, respectively). Baseline performance for consonants /th/, /f/, and /p/ were low enough to allow demonstration of improvement. The baseline score for /th/ was 42%, for /f/ was 59%, and for /p/ was 79%. Of these three consonants, a large improvement of 33% was obtained for the low intensity fricative /f/. This is in good agreement with the 38% improvement shown by Kennedy. The improvements for /p/ and /th/ were very small in the present study, while Kennedy reported approximately 22% improvement for the /p/ and approximately 12% for /th/.

The CE Peak errors were compared with the error patterns in the similar condition in the Kennedy et al (1998) study. Substitution of /t/ for /p/ in the context of /a/ was 25% in both studies, while substitution of /f/ for /th/ was 74% in the Kennedy et al (1998) study and 83% in the current study.

In the context of /i/, a fairly close correspondence was noted for the /f/ for /th/ substitution between the two studies; however the /th/ for /f/ substitution was 26% in the current study, while it was only 12% in the Kennedy study. Similarly, the /t/ for /p/ substitution was 20% in the Kennedy et al study, while it was 32% in the current study.

CE Peak and COMP65 Results

Voiceless consonant recognition was 10% higher in the CE Peak condition than in the COMP65 condition. The COMP65 condition reflects the type of performance that the hearing-impaired listener will obtain for speech at average conversational level. While both methods are meant to improve the C-V ratio and both have the goal of increasing speech recognition, there are fundamental differences in the manner in which C-V ratio enhancement is obtained in the two different methods. The CE Peak condition is obtained by amplifying the consonant level without changing the vowel level. Compression causes change in both vowel and consonant levels. The fact that the CE Peak condition was found to yield better recognition of voiceless consonants suggests that it matters how C-V ratio alterations are obtained. Hickson, Thayer and Bates (1999) found that compression amplification altered C-V ratio more than linear amplification did, and this alteration varied as a function of compression ratio, crossover frequency and input level. It is possible that use of a different algorithm for setting the parameters of the compression hearing aid might not have changed the vowel levels quite as much and may have resulted in better consonant recognition.

The effect of compression on C-V ratio (65 dB input)

In the current study, C-V ratio increased with compression amplification. With the compression parameters of the two channel hearing aid set according to the NAL-NL1 prescriptive formula and with a 65 dB speech input (representative of average conversational speech level), all syllables showed increases in C-V ratio from the baseline condition. The C-V ratio in the COMP65 condition was found to change dramatically in different syllables. The change in C-V ratio from baseline ranged from -2 dB to 22 dB. The magnitude of the changes found in the present study is similar to that reported by Hickson and Byrne (1995).
The amount of C-V ratio enhancement obtained with compression was rarely equal to that found to be optimal using the CE Peak method. In most cases, The CE Peak condition had a higher C-V ratio, although for eight of the twenty-one syllables, the COMP65 condition had higher C-V ratios. The majority of these were in vowel /u/ context.

There were some cases where the C-V ratio did not change with compression, although increases in the amplitude of the consonant could have been beneficial. This was the case with the consonant /f/ in the vowel contexts of /i/ and /a/. With a 65 dB speech input, compression did not achieve a change in C-V ratio. In contrast, in the CE Peak condition the C-V ratio was increased from baseline by 15 dB yielding a substantial increase in recognition of 58%. Performance in the context of /a/ showed a similar pattern. The C-V ratio was not changed by compression and scores similar to baseline were obtained. The increases in C-V ratio in the CE Peak condition led to a 24% improvement in recognition (from baseline).

In the case of the syllable /ap/, the C-V ratio obtained with compression (65 dB input) yielded a similar C-V ratio to that obtained with CE Peak, but resulted in different recognition scores. The C-V ratio for /ap/ in the baseline condition was -28 dB. The C-V ratio for the COMP65 condition was -14 dB and for the CE Peak condition the C-V ratio was -16.59. Consonant recognition scores were 83%, 30%, and 95% for baseline, compression, and CE Peak conditions, respectively. Examination of the differences in the RMS levels of the vowel and consonant components of the three conditions may help to explain the large differences in recognition with such similar C-V ratios. The levels (as presented in the Results section) are with reference to full scale on the measurement software. The vowel level in the original syllable (baseline) was 71.69 dB. The vowel level in the COMP65 condition was 64.8 dB. In the CE Peak condition, the vowel level was 71.78 dB. Thus the vowel level in the baseline and CE Peak condition are the same, but the vowel level has been decreased by approximately 7 dB in the compression condition. The consonant level in the baseline condition was 43.7 dB. In the compression condition, the consonant level was 50.7 dB. In the CE Peak condition, the consonant level was 55.19 dB. Thus, in the compression condition the consonant level increased by 7 dB and in the CE Peak condition the consonant level increased by almost 12 dB. Thus, while the compression condition increased the C-V ratio by 14 dB, half of that increase was from decreasing the vowel and half was from increasing the consonant level. Clearly, this was not the same as increasing the consonant by 12 dB. The result of the compression was to cause degradation in performance, despite the increased C-V ratio.

Effect of Speech Input level on C-V ratio enhancement and speech recognition performance with Compression Amplification

Since amplitude compression is a non-linear form of amplification, the second aim was to determine the effect of input level on C-V ratio and on recognition of voiceless consonants. Acoustic analyses included measures of the levels of the vowel and consonants, C-V ratio, and wideband spectrograms of selected syllables. Consonant recognition was assessed using the same nonsense VC syllables containing voiceless consonants in combination with the vowels /a/, /u/, and /i/.

The main findings were:

1. Input level affected the C-V ratio obtained with compression amplification. C-V ratio increased with increasing input level.
2. On average, consonant recognition increased with increasing input level. There was, however, an interaction between compression input level and consonant. Recognition of consonants /s/ and /sh/ was similar for all three input levels. Recognition of consonants /t/, /k/, and /th/ was lower for the 50 dB input, but was similar for 65 and 80 dB inputs. Recognition of /f/ increased with increasing level. For the consonant /p/, the lowest performance was obtained at

the 50 dB input level, recognition at the 65 dB level was best, and recognition at the 80 dB level was poorest.

3. The pattern of errors revealed that the alveolar stop /t/ was substituted for /p/ in the compression conditions. The number of substitutions increased with increasing input level. A comparison of the spectral cues of the syllables containing /p/ and /t/ in the baseline (natural stimulus), compression, and CE Peak conditions revealed that compression amplification altered the acoustic cue in /p/ by increasing the amplitude of the spectral peak (burst) in the higher frequency regions relative to the amplitude of the vowel formant in the higher frequency regions. Because the amplitude of the spectral burst relative to the vowel formants serves as a cue to place of articulation, the increased amplitude of the burst in /p/ after compression caused the compressed /p/ to be perceived as /t/.

Changes in C-V ratio with changes in input level

Compression amplification increased the C-V ratio of most consonants for all input levels. On average, the highest C-V ratios were obtained with the highest input level. Average increases in C-V ratio were approximately 7dB with a 50 dB input, 9 dB with a 65 dB input, and 11 dB with an 80 dB input. Although it is not possible to make a direct comparison of the C-V ratios obtained in the present study to those obtained by Hickson and colleagues (1999) due to differences in the compression characteristics used in the two studies, the present results confirm the finding of Hickson that compression amplification does increase the C-V ratio and that higher C-V ratios will be obtained for higher input levels.

There was an interaction of the compression input level with the vowel environment. In the COMP50 condition, the largest C-V ratio increases occurred in /u/ context and both positive and negative changes in C-V ratio occurred in /a/ context. In the /i/ vowel environment, many of the changes were in the negative direction. The few increases in C-V ratio with /i/ were small. In the COMP65 condition, again the largest increases in C-V ratio were in the context of /u/, followed by /a/. Improvements in the context of /i/ were minimal. Interestingly, the vowel levels in COMP65 were lower, compared to the Baseline condition, while the consonant levels increased. The same pattern of changes occurred in COMP80 condition, except that C-V ratios increases were larger in the /i/ context for this input level than at other input levels.

Consonant recognition as a function of input level

The finding that consonant recognition increased with increasing level is not surprising. On average, the poorest performance was in the COMP50 condition and best in the COMP80 condition. Performance was poor in all three compression conditions for /p/ and /th/ and performance under all three compression conditions was fairly equal for /s/ and /sh. Consonant recognition with COMP80 was significantly higher than the other two compression conditions for all consonants except /p/ and /th/. The recognition improved from the 50 to the 65 dB input, but then decreased with the 80 dB input. This "rollover" was seen for /p/ in the context of all three vowels. Rollover in the COMP80 condition may be due to the higher output level or to higher C-V ratios in the COMP80 condition. The consonant and vowel levels were higher in the COMP80 condition for /p/ in the context of all three vowels than in any other condition (including CE Peak).

Pattern of errors

Examination of the error patterns revealed that the most common substitutions in the baseline and CE Peak conditions occurred for weak consonant sounds /th/ and /f/. In the context of vowel / a/, substitution of /f/ for /th/ was 70% in the baseline condition and 84% in the CE Peak condition; while /th/ for /f/ was 24% for both baseline and CE Peak.

The majority of errors were made in the compression conditions. With a 50 dB speech input, substitutions were distributed more across consonants than with the higher input levels. It is possible in this condition that lack of audibility of the weaker consonants may have played a role. For the target sound /k/, errors were found only in the COMP50 condition where the sound was substituted with /f/ (24%) and /p/ (23%). Substitutions for the target sound /f/ included /p/ (30%) and /th/ (18%). Substitutions for the target /p/ included /t/ (48%), /f/ (15%) and /th/ (16%). Thus low intensity sounds were substituted for other low intensity sounds.

With 65 dB input, the error patterns for /f/ and /th/ were more similar to baseline and CE Peak; i.e., confusion between /f/ and /th/ in both directions. Substitution of /f/ for /th/ was 65%, and /th/ for /f/ was 41%. With the 80 dB input,

/f/ was also substituted for /th/ (81%), but the /th / for /f/ substitution was less than 15%.

The substitution of /t/ for /p/ was observed only in the compression conditions. The largest percentage of /p/ for /t/ substitution was under COMP80 (51%) followed by COMP65 (40%), CE Peak (29%) and Baseline (19%). After the study was completed, the experimenter and several listeners with normal hearing listened to the compressed stimuli containing /p/ and all agreed that the syllables sounded as if they contained the /t/ sound.

In order to determine why the compression of stimuli containing /p/ were perceived as /t/, spectrographic analyses were carried out on all syllables containing /p/ and /t/. A comparison of these analyses revealed that after compression, the amplitude of the burst of the consonant /p/ was greater than that of the formant of the preceding vowel and thus was more similar to what was seen in the syllables with /t/. This increase in amplitude is probably responsible for the distortion resulting in /t/ for /p/ confusion. The cues for perception of /p/ and /t/ have been studied by others (e.g., Ohde and Stevens, 1983; Hedrick, Schulte and Jedsteadt, 1995; Hedrick and Rice, 2000). That research has shown that in CV context, when the amplitude of the consonant burst is higher than that of the adjacent vowel in the F4 region, listeners responses tended to be in the direction of alveolar (/t/) than labial (/p/). It appears that because compression caused changes to both the vowel and consonant levels, the relative amplitude cues used to identify the bilabial voiceless consonant were changed sufficiently to result in the perception of an alveolar voiceless consonant. It is interesting that manipulation of the consonant amplitude in the CE Peak condition did not yield the same problem, probably because the vowel amplitude was not modified.

Implications of findings for clinical practice

Wide dynamic range syllabic compression of the type used in this study is becoming one of the most common types of amplification. At the outset of the experiment it was hypothesized that this form of compression would yield similar increases in recognition of voiceless consonants as direct C-V ratio enhancement. This was not the case with the compression parameters used in the two channel digital hearing aid tested in this study. It is possible that use of different compression parameters might have yielded different results. As mentioned earlier, parameters of the hearing aid such as the number of channels available in the hearing aid, the crossover frequencies, the compression ratios, and the attack release times will all affect the C-V ratio of the speech processed through the compression amplifier. Thus both the form of compression used, as well as the prescriptive formula will affect outcomes. It is also important to note that only a small number of consonants were tested in three vowel environments and that a single talker recorded the stimuli. Further investigations should consider different prescription methods, a wider range of test stimuli, and multiple talkers.

An additional factor to consider in the interpretation of findings of the present study is the fact that the bandwidth of the compression condition is narrower than that of the baseline and CE Peak conditions. The baseline condition represents a flat frequency broad band response without any filtering, and the CE Peak approximates a broad band frequency response with high frequency emphasis. The compression condition, on the other hand, has the narrower bandwidth typical of hearing aids (250-5000 Hz). However, because the participants all had severe hearing loss at frequencies above 4000 Hz, it is unlikely that the bandwidth limitation accounts for the lack of improvement in the compression condition. It is more likely that a fundamental difference in the manner of obtaining C-V ratio enhancement, as well as a need for optimizing the amount of C-V ratio enhancement account for differences between the COMP65 and CE Peak conditions.

Conclusions

In summary, the findings of this study confirm previous findings that modification of the consonant-vowel ratio using consonant enhancement technique can improve speech intelligibility. The anticipated increase in consonant perception from compression amplification was negligible and did not equal the gains obtained through consonant enhancement.

Compression amplification yielded marked improvements in C-V ratio compared to the un-processed condition, especially for the weaker consonants. However these C-V ratio improvements were often the result of decreasing vowel levels and increasing consonant levels. This is in contrast to the CE Peak condition, where only the consonant level was changed. Compression amplification resulted in higher C-V ratios than CE Peak in some of the compression conditions, but the highest C-V ratio did not necessarily result in the best consonant recognition. On average, consonant recognition under the compression conditions was poorer than in the CE Peak condition and quite often poorer than in the baseline condition.

Compression amplification disrupted spectral cues important to the recognition of the bilabial voiceless stop consonant in a manner that direct C-V ratio enhancement did not.

Appendix 1a-1b

Instructions and Procedures for establishing Most Comfortable Level of Loudness and Uncomfortable Level of Loudness

Appendix 1a

Procedure for establishing most comfortable level of loudness:

A lower level MCL and upper level MCL were obtained and the mid–point of the two values were set as the MCL.

Lower limit MCL procedure:

1. Listeners were instructed on lower limit MCL procedure.
2. A recorded loop of the carrier phrase “you will mark” was increased in 10 dB steps from the listener’s Speech Reception Threshold (SRT) until the listener indicated that the speech was at a just comfortable level.
3. The attenuator setting for this level was noted.
4. Signal level was decreased by 10dB, and increased in 5-dB steps until the listener indicated that speech was at a just comfortable level. If this level was identical to the initial response level for just comfortable speech, signal level was decreased in 10dB steps and increased in 2-dB steps and Lower level MCL was set at the point where the same level was identified three times.

Instructions for Comfortable Level Test - Lower Level:

The purpose of this test is to find the softest level for comfortable listening. We want you to decide when the speech is at a level which you feel is at a soft but comfortable listening level. Raise your hand when the speech reaches a just comfortable listening level.

Instructions for Comfortable Level Test - Upper Level:

The purpose of this test is to find and maintain loudness at which speech is just below a loud level, yet comfortable to listen. We want you to decide when the speech is at the highest level that is comfortable for you to listen to. Raise your hand each time you hear the highest level for comfortable listening.

Upper limit MCL procedure:

1. Listeners were instructed on upper limit UCL procedure.

2. A recorded loop of the carrier phrase “you will mark” was increased in 10 dB steps from the listener’s lower limit MCL until the listener indicated that the speech was louder than Most Comfortable Level. Attenuator setting at this level was noted.

3. Signal level was decreased by 10dB and increased in 2-dB steps until listener indicated that speech was at a high, yet comfortable level. If this level was identical to the initial response level for high level of comfort upper MCL was set at the point where the same level was identified three times.

Final MCL:

Mid-point between the lower and upper MCL was the MCL at which the test stimuli was presented.

Appendix 1b

Uncomfortable loudness level procedure:

1. Testing was done with each syllable presented at MCL through a COMPAQ ARMADA E 500 laptop computer.
2. The intensity of the consonant in each VC was enhanced in 3-dB steps until the listener indicated that the level was uncomfortable. The listener indicated the level at which the stimuli was uncomfortable using hand raise.
3. Stimuli were reduced by 6-dB and step 2 was repeated. UCL was set at the level at which UCL was identified at the same level of enhancement three times.
4. If a listener reached 24-dB of enhancement and failed to cross UCL, stimuli were repeated at this level once, and if there was no change, stimuli were presented at the maximum level of enhancement (24 dB).

Appendix 2a-2c

RMS values for each Vowel and each Consonant of the voiceless syllables, corresponding C-V ratios and Consonant recognition Scores for all five experimental conditions.

Appendix 2a RMS values for V-C, C-V ratios and Corresponding Scores for /a/

V/C	/a/	/p/	CVR	SCORE
BSL	71.69	43.69	-28	82%
COMP50	48.78	32.54	-16.24	14%
COMP65	64.8	50.7	-14.1	30%
COMP80	75.67	64.51	-11.16	18%
CEPEAK	71.78	55.19	-16.59	95%
V/C	/a/	/t/	CVR	SCORE
BSL	69.6	45.44	-24.16	89%
COMP50	48.75	32.61	-16.14	58%
COMP65	63.12	49.73	-13.39	96%
COMP80	74.17	62.84	-11.33	96%
CEPEAK	70.03	58.12	-11.91	98%
V/C	/a/	/k/	CVR	SCORE
BSL	68.05	41.65	-25.73	98%
COMP50	47.13	20.67	-26.46	73%
COMP65	61.32	41.68	-19.64	95%
COMP80	72.5	56.62	-15.88	95%
CEPEAK	70.16	55.4	-14.76	99%
V/C	/a/	/f/	CVR	SCORE
BSL	70.8	42.28	-28.42	67%
COMP50	50.46	21.18	-29.28	43%
COMP65	64.31	41.86	-22.45	49%
COMP80	75.23	55.49	-19.74	84%
CEPEAK	69.52	52.22	-17.3	90%
V/C	/a/	/s/	CVR	SCORE
BSL	71.13	59.64	-11.49	99%
COMP50	50.52	44.45	-6.07	98%
COMP65	64.16	60.3	-3.86	96%
COMP80	75.02	72.13	-2.89	98%
CEPEAK	70.79	69.07	-1.72	99%
V/C	/a/	/sh/	CVR	SCORE
BSL	68.46	57.81	-10.65	96%
COMP50	47.83	45.89	-1.94	92%
COMP65	62.48	61.46	-1.02	96%
COMP80	73.82	73.21	-0.61	95%
CEPEAK	69.96	68.77	-1.19	95%
V/C	/a/	/th/	CVR	SCORE
BSL	70.08	46.56	-23.52	23%
COMP50	49.49	21.72	-27.77	21%
COMP65	63.6	42.17	-21.43	34%
COMP80	74.49	57.2	-17.29	13%
CEPEAK	70.62	54.42	-16.2	20%

Appendix 2b. RMS values for V-C, C-V ratios and Corresponding Scores for /u/

V/C	/u/	/p/	CVR	SCORE
BSL	70.79	45.93	-24.86	82%
C 50 dB	37.18	29.7	-7.48	35%
C 65 dB	55.01	49.55	-5.46	49%
C 80 dB	68.36	63.58	-4.78	33%
CEPEAK	68.07	54.43	-13.64	69%
V/C	/u/	/t/	CVR	SCORE
BSL	70.26	49.33	-20.93	97%
C 50 dB	39.26	35.63	-3.63	91%
C 65 dB	55.66	53.21	-2.45	98%
C 80 dB	68.42	66.14	-2.28	98%
CEPEAK	68.08	58.56	-9.52	99%
V/C	/u/	/k/	CVR	SCORE
BSL	70.65	46.07	-24.58	96%
C 50 dB	34.43	23.12	-11.31	31%
C 65 dB	52.44	43.36	-9.08	94%
C 80 dB	65.45	57.44	-8.01	98%
CEPEAK	68.41	53.43	-14.98	99%
V/C	/u/	/f/	CVR	SCORE
BSL	68.83	47.8	-21.03	75%
C 50 dB	31.02	26.65	-4.37	30%
C 65 dB	49.92	45.71	-4.21	85%
C 80 dB	63.79	59.98	-3.82	95%
CEPEAK	66.98	57.1	-9.88	97%
V/C	/u/	/s/	CVR	SCORE
BSL	69.44	61.76	-7.68	98%
C 50 dB	34.72	48.16	13.44	96%
C 65 dB	52.85	63.34	10.49	99%
C 80 dB	66.58	74.52	7.94	98%
CEPEAK	67.11	70.68	3.57	99%
V/C	/u/	/sh/	CVR	SCORE
BSL	68.24	56.04	-12.11	98%
C 50 dB	32.32	44.21	11.89	97%
C 65 dB	51.36	60.41	9.05	99%
C 80 dB	65.12	72.11	6.99	98%
CEPEAK	66.49	65.2	-1.29	99%
V/C	/u/	/th/	CVR	SCORE
BSL	68.98	45.54	-23.33	80%
C 50 dB	32.25	22.5	-9.75	21%
C 65 dB	50.88	42.89	-7.99	90%
C 80 dB	64.6	57.34	-7.26	96%
CEPEAK	66.36	53.92	-12.44	92%

Appendix 2c- RMS values for V-C, C-V ratios and Corresponding Scores for /i/

V/C	/i/	/p/	CVR	SCORE
BSL	68.44	46.07	-22.37	72%
C 50 dB	56.38	29.8	-26.58	34%
C 65 dB	70.47	49.17	-21.3	47%
C 80 dB	78.74	63.1	-15.64	15%
CEPEAK	67.72	53.56	-14.16	70%
V/C	/i/	/t/	CVR	SCORE
BSL	66.87	52.65	-14.22	98%
C 50 dB	51.14	41.55	-9.59	92%
C 65 dB	66.01	57.32	-8.69	98%
C 80 dB	76.2	68.58	-7.62	98%
CEPEAK	66.48	64.03	-2.45	97%
V/C	/i/	/k/	CVR	SCORE
BSL	67.37	51.9	-15.47	99%
C 50 dB	51.49	39.23	-12.26	95%
C 65 dB	66.58	55.43	-11.15	99%
C 80 dB	76.84	68.03	-8.81	99%
CEPEAK	66.74	64.27	-2.47	97%
V/C	/i/	/f/	CVR	SCORE
BSL	66.81	45.24	-21.57	34%
C 50 dB	52.51	27.72	-24.79	47%
C 65 dB	67.52	46.24	-21.28	28%
C 80 dB	77.35	59.91	-17.44	47%
CEPEAK	67.29	61.25	-6.04	92%
V/C	/i/	/s/	CVR	SCORE
BSL	66.42	64.36	-2.06	97%
C 50 dB	51.01	45.67	-5.34	96%
C 65 dB	66.24	61.43	-4.81	99%
C 80 dB	77.03	72.89	-4.14	98%
CEPEAK	64.44	74	9.56	98%
V/C	/i/	/sh/	CVR	SCORE
BSL	69.58	59.43	-10.15	97%
C 50 dB	55.57	47.52	-8.05	95%
C 65 dB	69.69	62.74	-6.95	96%
C 80 dB	78.21	73.79	-4.42	98%
CEPEAK	68.93	70.57	1.64	97%
V/C	/i/	/th/	CVR	SCORE
BSL	66.1	46.82	-19.28	22%
C 50 dB	51.84	26.23	-25.61	24%
C 65 dB	66.63	45.68	-20.95	14%
C 80 dB	77.5	60.54	-16.96	17%
CEPEAK	65.05	56.3	-8.75	18%

Appendix 3a-3e
Confusion Matrices for all five experimental conditions

Appendix 3a. Confusion martrix for Baseline condition

BASELINE	af	ak	ap	as	ash	at	ath
af	144	1	36	0	0	1	58
ak	0	238	0	0	0	0	2
ap	18	0	183	0	0	37	2
as	0	0	0	240	0	0	0
ash	0	0	0	0	240	0	0
at	16	0	1	1	0	202	20
ath	167	2	10	1	0	2	58

BASELINE	if	ik	ip	is	ish	it	ith
if	73	0	2	1	0	4	160
ik	0	239	0	0	0	0	1
ip	0	0	147	0	0	73	20
is	1	1	0	231	1	1	5
ish	0	0	0	0	239	1	0
it	0	1	0	0	0	228	11
ith	150	0	1	0	20	1	68

BASELINE	uf	uk	up	us	ush	ut	uth
uf	171	1	28	0	0	3	37
uk	0	224	2	0	0	8	6
up	4	1	179	0	0	46	10
us	1	0	0	234	4	0	1
ush	0	0	0	1	238	0	1
ut	1	0	0	0	0	228	11
uth	44	1	4	0	0	16	175

Appendix 3b. Confusion matrices for COMP50.

COMP50	af	ak	ap	as	ash	at	ath
af50	91	11	71	0	1	23	43
ak50	15	168	8	5	0	15	29
ap50	37	3	44	2	0	116	38
as50	3	0	0	234	0	0	3
ash50	0	0	0	0	240	0	0
at50	32	11	8	8	0	137	44
ath50	80	13	55	7	1	42	42

COMP50	if	ik	ip	is	ish	it	ith
if50	96	5	24	7	0	30	78
ik50	0	240	0	0	0	0	0
ip50	52	3	86	4	0	46	49
is50	1	0	0	227	3	0	9
ish50	0	0	0	0	240	0	0
it50	3	1	0	9	0	208	19
ith50	67	14	50	7	0	35	67

COMP50	uf	uk	up	us	ush	ut	uth
uf50	55	17	33	1	0	85	49
uk50	57	83	56	1	0	17	26
up50	39	30	93	1	0	48	29
us50	0	0	0	228	11	1	0
ush50	0	0	0	1	239	0	0
ut50	11	0	0	2	1	210	16
uth50	90	17	29	1	0	34	69

Appendix 3c. Confusion matrices for COMP65.

COMP65	af	ak	ap	as	ash	at	ath
af65	113	2	20	0	0	7	98
ak65	0	220	0	0	0	15	5
ap65	0	0	85	0	0	147	8
as65	0	0	0	240	0	0	0
ash65	0	0	0	0	240	0	0
at65	0	0	1	0	0	225	14
ath65	155	1	3	1	1	0	79

COMP65	if	ik	ip	is	ish	it	ith
if65	58	0	4	4	0	6	168
ik65	0	240	0	0	0	0	0
ip65	0	1	114	0	0	109	16
is65	0	0	0	240	0	0	0
ish65	0	0	0	0	240	0	0
it65	0	0	0	0	0	237	3
ith65	198	0	0	0	0	0	42

COMP65	uf	uk	up	us	ush	ut	uth
uf65	193	0	0	0	0	0	47
uk65	6	216	3	0	0	1	14
up65	0	8	116	0	0	95	21
us65	0	0	0	238	2	0	0
ush65	0	0	0	0	240	0	0
ut65	0	0	0	0	0	234	6
uth65	27	0	1	1	0	0	211

Appendix 3d. Confusion matrices for COMP80.

COMP80	af	ak	ap	as	ash	at	ath
af80	184	0	0	29	0	0	27
ak80	3	220	0	0	0	15	2
ap80	0	1	56	0	0	170	13
as80	0	0	0	238	2	0	0
ash80	0	0	1	0	239	0	0
at80	2	0	0	0	0	226	12
ath80	195	0	0	10	4	0	31

COMP80	if	ik	ip	is	ish	it	ith
if80	89	0	0	16	0	0	135
ik80	0	240	0	0	0	0	0
ip80	0	1	50	0	0	183	6
is80	0	0	0	238	1	0	1
ish80	0	0	0	0	240	0	0
it80	0	0	0	0	0	234	6
ith80	157	0	0	1	18	0	64

COMP80	uf	uk	up	us	ush	ut	uth
uf80	216	0	0	0	0	0	24
uk80	1	235	0	0	0	3	1
up80	0	20	90	0	0	123	7
us80	0	0	0	236	4	0	0
ush80	0	0	0	2	238	0	0
ut80	0	0	0	0	0	237	3
uth80	1	0	0	3	0	1	235

Appendix 3e. Confusion Matrices for CE Peak

CEPeak	af	ak	ap	as	ash	at	ath
af	177	0	0	1	0	4	58
ak	1	239	0	0	0	0	0
ap	0	0	180	0	0	60	0
as	0	0	0	239	0	1	0
ash	0	0	0	0	240	0	0
at	0	0	0	0	0	231	9
ath	201	0	0	2	0	1	36

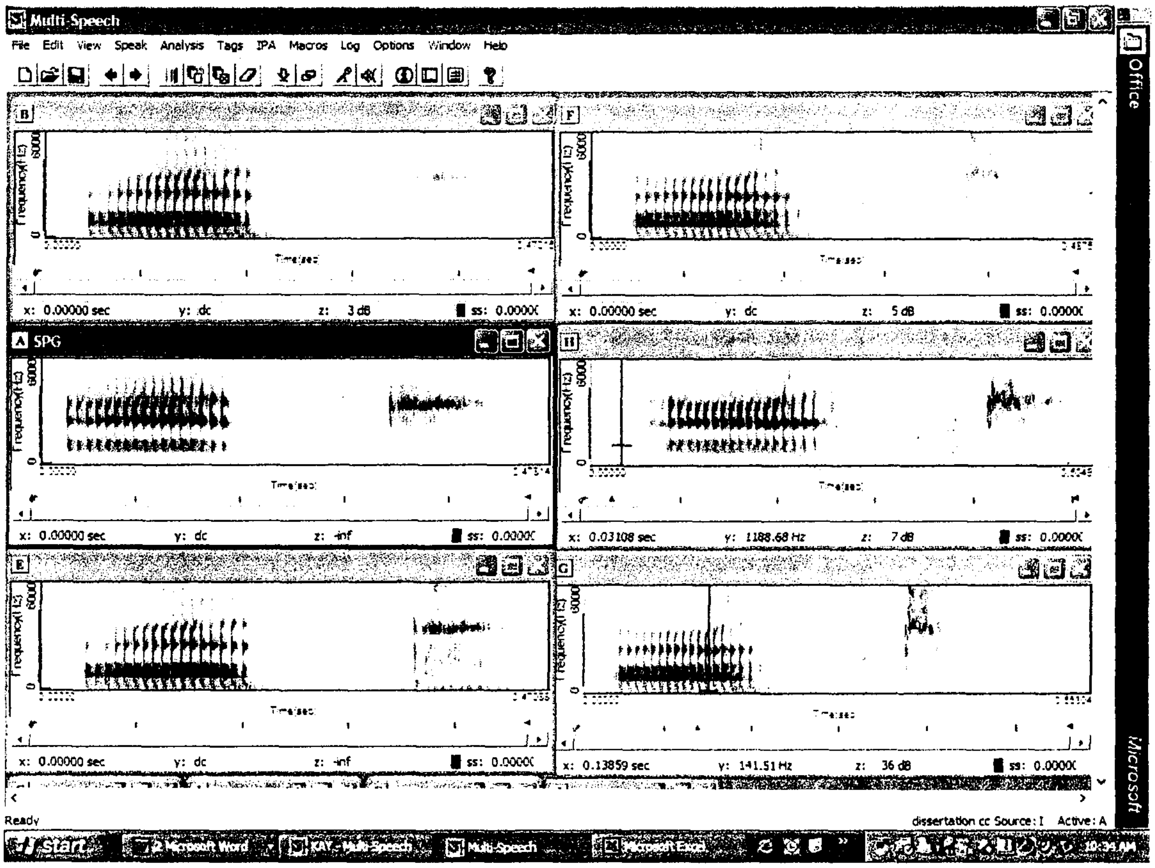
CEPeak	if	ik	ip	is	ish	it	ith
if	177	0	0	0	0	0	63
ik	2	227	1	0	0	10	0
ip	0	1	155	0	0	76	8
is	0	0	1	234	0	0	5
ish	2	0	0	0	238	0	0
it	1	0	0	0	0	227	12
ith	172	0	0	0	12	0	56

CEPeak	uf	uk	up	us	ush	ut	uth
uf	224	0	0	0	0	0	16
uk	0	239	0	0	0	1	0
up	1	10	153	0	0	69	7
us	0	0	0	234	3	0	3
ush	0	0	0	0	240	0	0
ut	0	0	0	0	0	238	2
uth	39	0	0	0	0	0	201

Appendix 4a-4c

Spectrograms of /p/ and /t/ in th contexts of /a/, /u/ and /i/

Appendix 4a. Spectrograms of /ap/ (left panels) and /at/ (right panels) in the Baseline (top), COMP80 (middle) and CE Peak (bottom) conditions.



x: 0.00000 sec y: dc z: 3 dB ss: 0.0000(

x: 0.00000 sec y: dc z: 5 dB ss: 0.0000(

x: 0.00000 sec y: dc z: -inf ss: 0.0000(

x: 0.03108 sec y: 1188.68 Hz z: 7 dB ss: 0.0000(

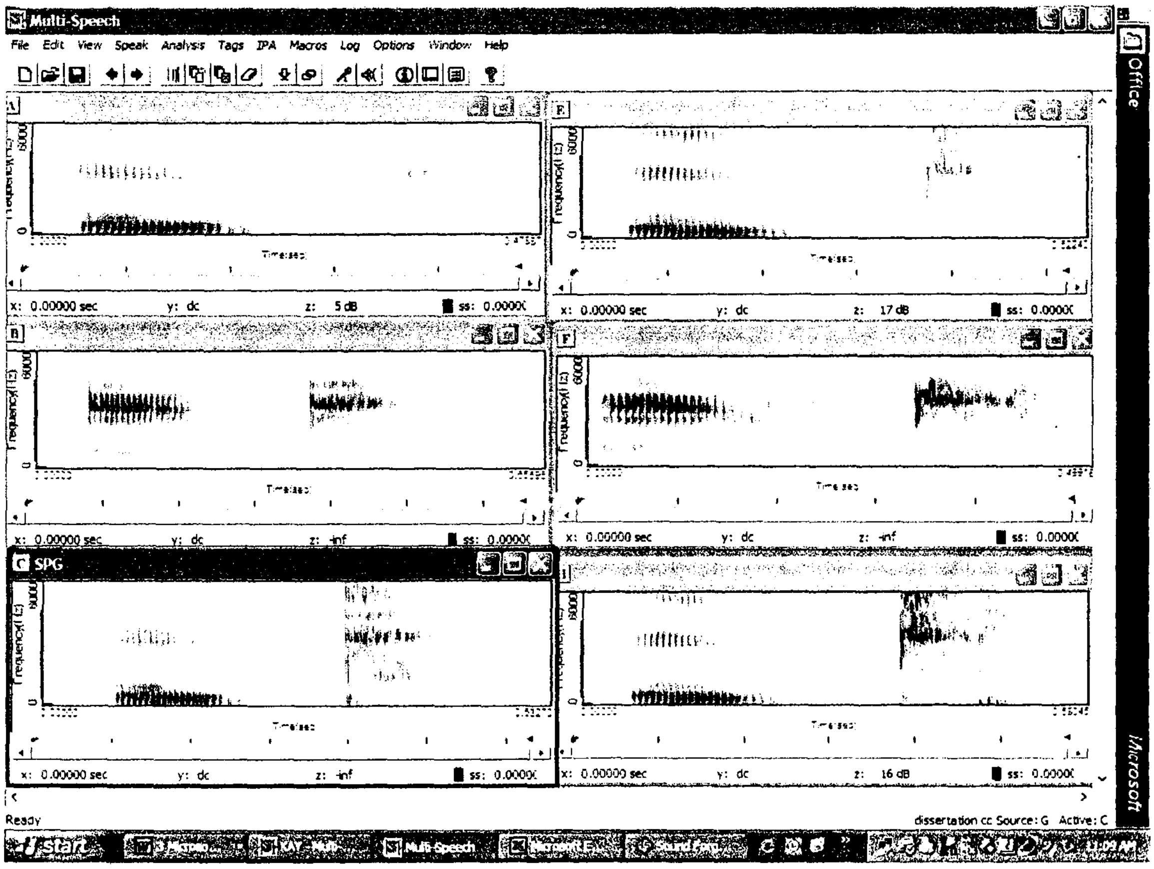
x: 0.00000 sec y: dc z: -inf ss: 0.0000(

x: 0.13859 sec y: 141.51 Hz z: 36 dB ss: 0.0000(

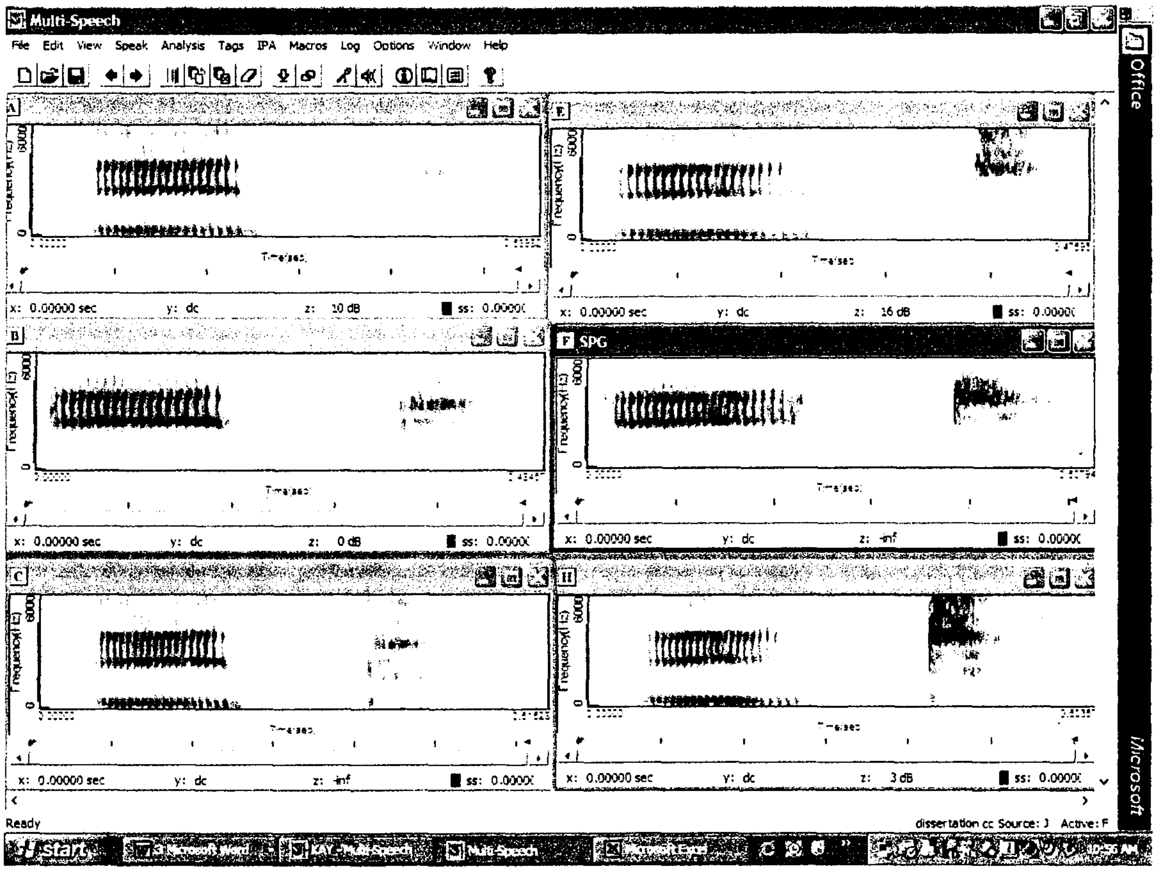
Ready

dissertation cc Source: I Active: A

Appendix 4b. Spectrograms of /up/ (left panels) and /ut/ (right panels) in the Baseline (top), COMP80 (middle) and CE Peak (bottom) conditions.



Appendix 4c. Spectrograms of /ip/ (left panels) and /it/ (right panels) in the Baseline (top), COMP80 (middle) and CE Peak (bottom) conditions



REFERENCES

Balakrishnan, U., Freyman, RL., Chiah, Y.C., Nerbonne, G. P. and Shea, K. J. (1996). Consonant recognition for spectrally degraded speech as a function of consonant-vowel intensity ratio. *J Acoust Soc Am* 99:3758-3769.

Braida, L. D., Durlach, N. I., DeGennaro, S. V., Peterson, P. M. and Bustamante, D. K. Review of recent research on multiband amplitude compression for the hearing impaired. Vanderbilt hearing aid report, UpperDarby, PA. Monographs in contemporary audiology, 1982: 133-140

Braida, L. D., Durlach, N. I., Lippmann, R. P., Hicks, B. L., Rabinowitz, W. M., and Reed, C. (1979) Hearing aids-A review of past research on linear amplification, amplitude compression, and frequency lowering. ASHA Monographs, 19.

Byrne, D., & Dillon, H. (1986). The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear and Hearing 7: 257-265.

Byrne, D., Dillon, H., Ching, T., Katsch, R., & Keidser, G. (2001). The NAL-NL1 procedure for fitting non-linear hearing aids: Characteristics and comparisons with other procedures. Journal of American Academy of Audiology, 12: 37-51.

Dillon, H. (2001) Advanced Signal Processing Schemes. In Dillon (Ed.) Hearing Aids. pp.205. Sydney; Boomerang Press

Dubno, J. & Levitt, H. (1981). Predicting consonant confusions from Acoustic Analysis, JASA 69, No. 1:249-261.

Freyman, R. L. and Nerbonne, G. P. (1989). The importance of consonant-vowel intensity ratio in the intelligibility of voiceless consonants. *J Speech Hear Disord* 32:524-535.

Gordon-Salant, S. (1986). Recognition of natural and time/intensity altered CV's by young and elderly subjects with normal hearing. *J Acoust Soc Am* 80:1599-1607.

Gordon-Salant, S. (1987). Effects of acoustic modification on consonant recognition by elderly hearing-impaired subjects hearing. *J Acoust Soc Am* 81:1199-1202.

Hecker, M. H. L. (1974). A study of the relationship between consonant and vowel ratios and speaker intelligibility. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Hedrick, M. S. & Jesteadt, W. (1996). Effect of relative amplitude, presentation level, and vowel duration on perception of voiceless stop consonants by normal and hearing-impaired listeners. . *J Acoust Soc Am* 100:3396-3407.

Hickson, L. and Byrne, D. (1995). Acoustic analysis of speech through a hearing aid: effects of linear vs compression amplification. *Aust J Audiol* 17(1):1-13.

Hickson, L., Dodd, B. and Byrne, D. (1995). Consonant perception with linear and compression amplification. *Scand Audiol* 24:175-184.

Hickson, L., Thayer, N. & Bates, D. (1999). Acoustic analysis of speech through a hearing aid: Consonant-vowel ratio effects with two-channel compression amplification, *J Am Acad Audiol* 10:549-556.

Kennedy, E., Levitt, H., Neuman, A. C and Weiss, M. (1998). Consonant-vowel intensity ratios for maximizing consonant recognition by hearing-impaired listeners. *J Acoust Soc Am* 103:1098-1114.

Levitt, H. (1982) Speech discrimination in the hearing impaired: Spectrum considerations. The Vanderbilt Hearing-Aid Report; State-of-the-Art Research Needs. Eds. Studebaker, G and Bess, F.Monographs in Contemporary Audiology

Montgomery, A. and Edge, R. (1988). Evaluation of two speech enhancement techniques to improve intelligibility for hearing-impaired adults. *J Acoust Soc Am* 31:386-393.

Ohde, R. N. & Stevens, K. N. (1983). Effect of burst amplitude on the perception of stop consonant place of articulation. . *J Acoust Soc Am* 74:706-714.

Preves, D., Fortune, T., Woodruff, B. & Newton, J. (1991). Strategies for enhancing the consonant-to-vowel intensity ratio with in-the-ear hearing aids. *Ear and Hearing, 12* (Suppl.), 139S-153S.

Sammeth, C. A., Tetzeli, M. and Ochs, M. T. (1996). Consonant recognition performance of hearing-impaired listeners using one linear and three nonlinear hearing aids. *J Am Acad Audiol* 7:240-250.

Smith, L. and Levitt, H. (1999) Consonant Enhancement effects on speech recognition. J Am Acad Audiol 10, 8:411-421

Villchur, E. (1973). Signal processing to improve speech intelligibility in perceptive deafness. J. Acous. Soc. Am. 53, 1646-1647.

Winer, B. J. Statistical Principles in Experimental Design, (1970). McGraw Hill, N.Y. 2nd Edition.