INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

THE EFFECTS OF HIGH-FREQUENCY EMPHASIS AND AMPLITUDE COMPRESSION ON THE SHORT-TERM INTENSITY RANGE OF SPEECH

by

BALAJI ORUGANTI

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

2000

UMI Number: 9969714

Bell & Howell 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

3-21-2000
Date

Arthur Boothroyd, Ph.D.
Chair of Examination Committee

3/28/00
Date

Richard G. Schwartz, Ph.D.
Executive Officer

Supervisory Committee:

Stanley Gelfand, Ph.D.
Queens College

Harry Levitt, Ph.D.
Graduate Center of CUNY

External Examiner:

Mark Ross, Ph.D.
League for the Hard of Hearing

City University of New York

ABSTRACT

THE EFFECTS OF HIGH-FREQUENCY EMPHASIS AND AMPLITUDE COMPRESSION ON THE SHORT-TERM INTENSITY RANGE OF SPEECH

By

Balaji Oruganti

Adviser: Professor Arthur Boothroyd

This study evaluated the effectiveness of simple high-frequency emphasis, and single-band and 2-band fast-release wide dynamic range amplitude compression in reducing the short-term dynamic range of speech. In subjects with severe and profound sensori-neural hearing loss, the dynamic range of hearing is often less than the short-term dynamic range of speech. To a certain extent the dynamic range of speech could be reduced by emphasizing the high-frequencies. But to go beyond this requires syllabic compression. The results of past research on syllabic compression have been equivocal. Conflicting results have been attributed to inappropriate selection of compression parameters, deleterious spectral and temporal changes accompanying compression, and evaluation methods.

The present study focused on the evaluation methods used to assess the effectiveness of compression. In the present study, the effectiveness of amplitude compression and high-frequency emphasis, and their benefits were measured both acoustically and perceptually. Single-band and 2-band syllabic compression and high-frequency emphasis of CVC words were effected through off-line digital processing. The acoustic evaluation included the measurement RMS cumulative amplitude distributions and dynamic input/output functions of the test stimuli. Perceptual measurement consisted of Performance-Intensity (PI) functions which were obtained for normally hearing pink-noise-masked individuals. The effects of signal processing on acoustic amplitude distribution, input/output function, and slope and width of the perceptual PI functions were evaluated. Results showed that, acoustically, the target compression ratio was achieved provided the spectral and temporal characteristics of the acoustical analysis matched those of the processing algorithms. Otherwise, the actual amount of compression was less than that target. Perceptually, a simple high-frequency emphasis, while did not result in a decrease of the overall dynamic range of speech, resulted in enhanced perception for low and moderate sensation levels. Addition of single-band compression with a compression ratio of 3:1 resulted in a reduction of the overall dynamic range of the high-frequency emphasized words by a factor of about 1.5:1. A 2-band compression with a compression ratio of 3:1 in both bands was found to be only marginally superior over single-band compression. The outcome of this research project has potential for improving assessment techniques of compression hearing aids.

To my wife Bhuvana, my children Swathi and Shreya

and

my beloved parents

ACKNOWLEDGMENTS

This study is the culmination of eight years study and work. I owe many thanks to the people who have provided me with guidance and encouragement along the way.

First, I would like to thank my mentor and dissertation chairman, Professor Arthur Boothroyd, for his exemplary encouragement and patience. It would have been difficult to complete this dissertation without his guidance during each stage of this study.

I would like to extend sincere thanks to the late Professor Irving Hochberg for all his guidance, support and encouragement during the course of my studentship. Professor Hochberg was an influential mentor during the formative years of my Ph.D. program.

Professor Harry Levitt deserves special recognition for his insightful comments and helpful suggestions. His passionate teachings on the topic of amplitude compression in hearing aids helped solidify my own interests in the topic.

I am also grateful to Professor Stanley Gelfand for his input and guidance during the planning of this study and analysis of study data.

Many thanks to Gary Chant for his help with instrument set up and calibration. In addition, I would like to thank Kristin Paulet for her help in editing this paper.

While I cannot list their names here, I would like to thank all the subjects who participated in this study, without whose cooperation the studies could not have been completed.

Finally, I would like to thank my loving wife Bhuvana and my children, Swathi and Shreya, for always being there. Their love and commitment have always provided encouragement, even during the most challenging times of this study. Thank you.

TABLE OF CONTENTS

LISTS OF TABLES

LIST OF FIGURES

LIST OF APPENDIXES

CHAPTER I: INTRODUCTION

1.1: GENERAL GOAL:

The general purpose of this study was to determine the potential benefits of fast-release wide dynamic range compression amplification for persons with sensorineural hearing loss. More specifically, the goal was to measure the effectiveness of three processing schemes in reducing both the acoustic dynamic range and the perceptual dynamic range of speech.

1.2: BACKGROUND:

Compression amplification is a general term that is applied to any active signal-processing scheme in which the acoustic dynamic range of the output signal is less than that of the input signal.

For present purposes, the term *acoustic dynamic range of speech* is used in a general sense to refer to the amplitude difference between the most intense elements in a speech signal and the least intense, but still perceptually useful elements. When speech is produced by a single talker, with constant effort, and at a fixed distance, the acoustic dynamic range is generally taken to be about 30 dB - this being the amplitude difference between the strong vowels *(/u/* and /ɑ/) and the weak fricatives *(*/f/ and /θ/). Because these speech elements last only a short time, this measure is often referred to as the *short-term dynamic range of speech*.

Variations of talker, talker effort, and talker distance introduce changes of average level of the speech signal. These variations can be as high as 20 dB or more, and they add to the overall dynamic range, resulting in a *long-term* *dynamic range of speech* of 50 dB or more. For example, the amplitude of the strong vowels in a listener's own speech can easily be 50 dB higher than that of the weak fricatives in the speech of another talker at a distance of 1 or 2 meters.

For the individual with normal hearing, the *dynamic range of hearing*, from threshold of audibility to threshold of uncomfortable loudness, is roughly 100 dB. This range is more than enough to accommodate the long-term dynamic range of speech. Persons with sensorineural hearing loss, however, typically exhibit elevated thresholds of audibility without corresponding elevations of the thresholds of uncomfortable loudness. In other words, the dynamic range of hearing is reduced compared with that of normally hearing individuals. As the degree of hearing loss increases, the dynamic range of hearing decreases. Moreover, within that range, the loudness increases more rapidly with increasing intensity than it does for persons with normal hearing. This phenomenon is commonly referred to as *loudness recruitment*.

The dynamic range of hearing of persons with severe and profound hearing loss is often less than the dynamic intensity range of speech. This situation creates a classical problem in hearing aid fitting. If gain is adjusted so that the most intense elements in the wearer's own speech remain below the threshold of discomfort, the least intense sounds of other people's speech become inaudible. On the other hand, if the gain is increased to provide full audibility of the speech of other people, the most intense sounds of the wearer's own speech will be intolerably loud. For persons with very severe or profound loss, the dynamic range of hearing can even fall below the short-term dynamic range of speech. In such cases, the hearing aid wearer cannot even be provided with audibility of the weaker elements in his or her own speech without causing discomfort from the more intense elements.

It is a customary practice in hearing aid fitting to limit the maximum output level to avoid discomfort. The simplest approach to output limiting is *peak clipping* in which the instantaneous output is prevented from exceeding a certain value. With properly adjusted peak clipping, it becomes possible to provide audibility of the weaker elements of speech without causing discomfort from the stronger elements. Unfortunately, peak clipping also introduces harmonic and inter-modulation distortions. Such distortions are generally believed to reduce hearing aid effectiveness, especially in noise.

As indicated earlier, compression amplification offers an alternative way to reduce the acoustic dynamic range of the speech signal, without the kinds of distortion caused by peak clipping. Because the output differs from the input, compression amplification is by no means free of distortion. It is generally believed, however, that the kinds of distortion introduced by compression amplification are less serious than those introduced by peak clipping.

The general principle behind compression amplification is that gain, instead of remaining fixed, is made dependent on the amplitude of the input signal. Based on their intended function, compression schemes can be broadly classified into three types:

1. Compression limiting: As its name implies, compression limiting is used to prevent the amplitude of the output from exceeding a certain value. Basically, gain remains constant until that limit is reached, at which point further increases of input are matched by an equal reduction of gain so that output becomes independent of input. Compression limiting systems typically have short release times so that gain can return quickly to its original value once the intense input has passed.

2. **Automatic gain control (AGC):** Although all compression systems involve automatic gain control, the term is usually used to refer to compression schemes with a long release time. When activated, AGC can eliminate or reduce long-term variations of average level that accompany changes of talker effort and distance, without disturbing the short-term dynamic range of speech. In some cases, the threshold at which compression is activated remains high, as in compression limiting, and average output for high level inputs is maintained at a constant level. In others, the threshold is lowered and, as input exceeds this threshold, the gain reduction is made less than the increase of input. The result is that the average level of the output continues to rise but at a lower rate than that of the input. The ratio of the two rates is known as the *compression ratio.*
3. **Wide dynamic range compression (WRDC):** As this term implies, WRDC is designed to operate over the full dynamic range of the speech input. In other words, the threshold is set so low that compression is activated by the weakest sounds in the input signal. A short release time is also typically used to reduce the variations of average speech level as well as the amplitude variations within the speech of a single talker at a fixed distance. This type of compression is becoming widely used as audiologists try to provide the hearing aid wearer with a

signal that remains audible, comfortable, and pleasant over a wide range of input levels. WRDC development is also being driven by the increased availability of sophisticated hearing aids incorporating powerful signal-processing capabilities.

The present study deals with the last type of processing namely fast-release wide dynamic range compression. Its effectiveness in reducing the short-term dynamic intensity range in the speech of a single talker with constant average level was evaluated.

Compression schemes can also be divided into single-band and multi-band systems. In a *single-band system*, overall gain is a function of overall signal level. Therefore, any adjustment of gain is applied equally across the frequency spectrum, making it impossible to change the intensity relationships among different parts of the spectrum. In a *multi-band system*, the signal is divided into two or more bands and gain adjustments in each band are a function of level in that band. With such a scheme, it becomes theoretically possible to change the intensity relationships among different parts of the spectrum. The present study involves comparison of a single-band scheme and a 2-band scheme.

Past research findings on the benefits of amplitude compression to the hearing-impaired have been equivocal. Whereas some data have shown compression amplification to be beneficial to these individuals, others have not. The negative findings have been attributed to the use of subjects who do not need compression; to the failure to match compression parameters to the characteristics of speech and/or the experimental subjects; to the deleterious spectral and temporal changes accompanying amplitude compression; to the acoustic outcome evaluation with steady-state stimuli rather than dynamic speech-like stimuli; and to perceptual outcome evaluation involving testing at high sensation levels at which compression offers little or no advantage.

In addition to these issues, there are two general weaknesses that can be identified in the research that has been carried out so far. First, there has been a failure to demonstrate that the compression schemes produce the desired acoustic results. The input/output characteristics of compression schemes are typically specified and measured in terms of their effects on steady-state signals. Speech, however, is far from a steady state. It contains rapid and dramatic changes of both spectrum and amplitude from moment to moment. The reductions of the dynamic range of speech may well be different from those predicted from performance with steady-state signals. This study included an attempt to measure the effectiveness of compression schemes in terms of their dynamic performance.

The second weakness has been a failure to demonstrate that the supposed perceptual benefits of compression are, in fact, present in the processed signal. Put another way, if hearing-impaired subjects fail to benefit from a compression scheme, one needs to know whether the problem lies with the characteristics of the hearing impairment or with inadequacies in the signal processing. The present study used normally hearing subjects to assess the effective reduction in the perceptual dynamic range of speech.

The *perceptual dynamic range of speech* is defined here as the increase of sensation level (SL) needed to go from bare audibility of the speech signal to near-perfect phoneme recognition.

In summary, compression amplification can reduce the dynamic range of speech making it theoretically possible for persons with very severe and profound hearing loss to attain full audibility of the weaker elements of the speech signal without experiencing discomfort from the more intense elements. Previous research on this topic has produced equivocal results. Among the weaknesses of previous research has been a failure to confirm that the compression schemes produce the intended effect on the speech signal both acoustically and perceptually. The present study was designed to measure the effectiveness of two compression schemes in terms of their ability to reduce both the acoustic dynamic range and the perceptual dynamic range of speech.

1.3: SPECIFIC GOALS:

1. To create a fast-release wide dynamic range compression algorithm with a 3:1 compression ratio in which the top 30 dB of the amplitude distribution of speech is compressed into a range of 10 dB.
2. To apply this algorithm in a single-band compression scheme and to compare the measured reduction of acoustic and perceptual dynamic ranges with the predicted values.
3. To apply this algorithm in a 2-band compression scheme and to test the hypothesis that a 2-band system is more effective than a single-band system in terms of producing the intended acoustic and perceptual outcomes.

1.4. OVERVIEW OF EXPERIMENTS:

The effectiveness of single-band and 2-band fast-release wide dynamic range compression schemes were compared with simple high-frequency-emphasis-alone and no-processing conditions. Thus, there were 4 conditions in this study. Both single-band and 2-band compression schemes were applied after pre-compression high-frequency emphasis.

Figure 1.1 illustrates the 4 test conditions. The dynamic range for the no-processing condition (CVC words) is about 30 dB. In general, the peak amplitude distribution of the low-frequency band is higher than the high-frequency band. In the second condition, the high-frequency band was amplified so as to equate the cumulative peak amplitude distributions of the low-frequency and high-frequency bands. In the third condition, a fast-release single-band wide dynamic range compression with a compression ratio of 3:1 was applied to the pre-compression high-frequency emphasized speech signal. In the fourth and final condition, a fast-release 2-band wide dynamic range compression with a compression ratio of 3:1 in both bands was applied to the pre-compression high-frequency emphasized speech signal. It was hypothesized that as we progress from the no-processing condition, to a high-frequency-emphasis-alone condition, to a pre-compression high-frequency emphasis plus single-band compression, to a pre-compression high-frequency emphasis plus 2-band compression condition, there would be progressively increased compression of the dynamic range of speech.

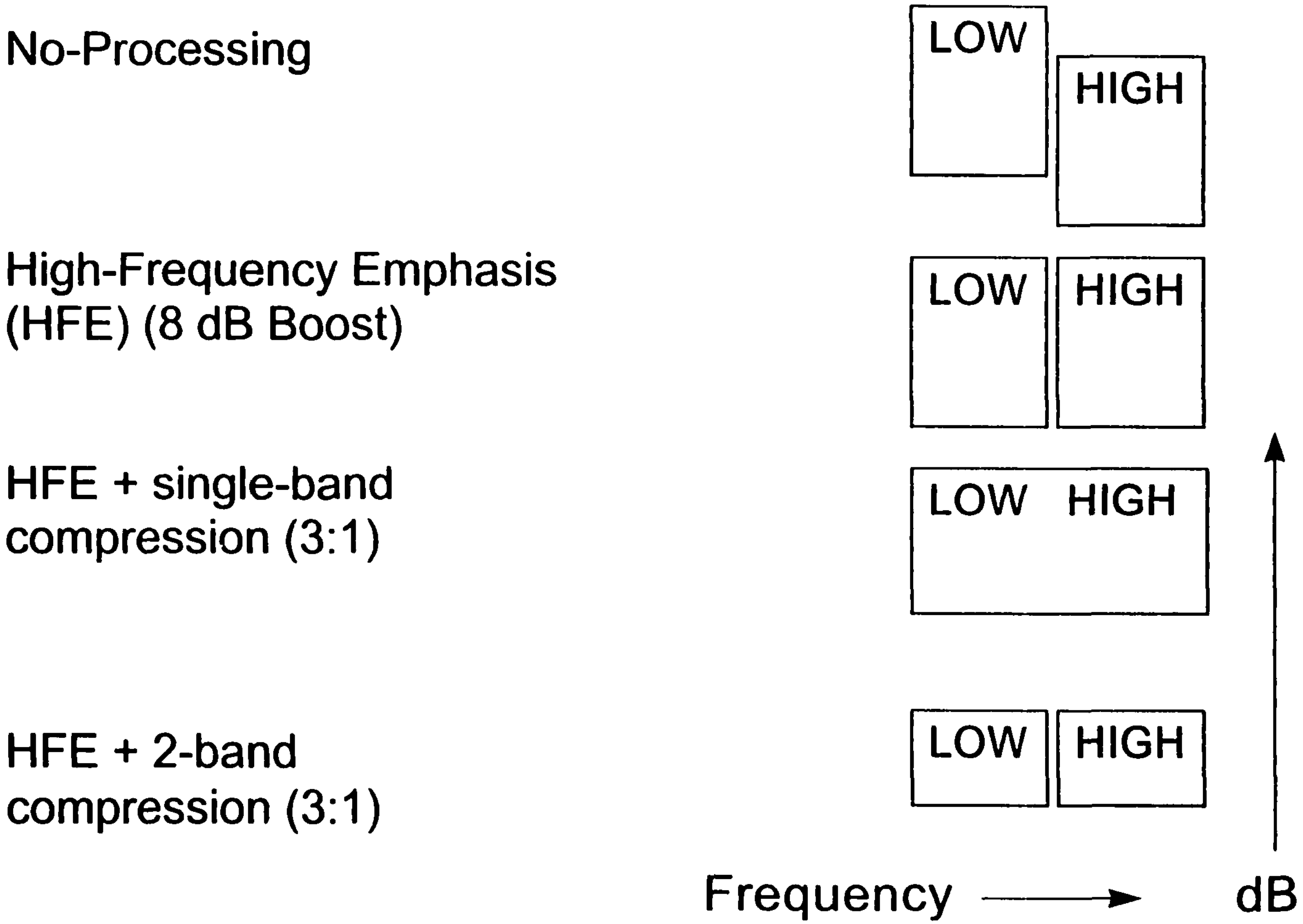

Figure 1.1: Schematic illustration of the dynamic intensity range of the 4 test conditions.

To evaluate the effectiveness of the different processing schemes, both acoustical confirmation and perceptual validation procedures, using speech stimuli, were included. The stimuli consisted of 20 iso-phonemic lists of consonant-vowel-consonant words. Each list consisted of 10 words. The same 10 vowels and 20 consonants appear in each list.

The acoustical evaluation consisted of measurements of cumulative RMS amplitude distributions and dynamic input/output functions. The input/output functions for the speech signal were obtained by plotting RMS amplitude distribution of the speech signal after compression against that before compression. The extent to which the obtained input/output function differed from a predicted function was measured to provide an estimate of the effectiveness of the compression schemes.

The perceptual outcome evaluation involved the measurement of performance-intensity (PI) functions in groups of pink-noise masked normal-hearing adults. The dependent variable was the percentage of correctly recognized phonemes. The relationship between the audibility of the speech signal and its corresponding effect on the PI function is illustrated in Figure 1.2. As the sensation level of the speech signal was raised from just above the threshold of audibility to about 30 dB SL, the percent phoneme recognition scores of naturally spoken CVC word stimuli increased from a guessing score of about 10% to a near-perfect score of about 90 to 100%. Thus, the perceptual dynamic range of naturally spoken CVC words was about 30 dB.

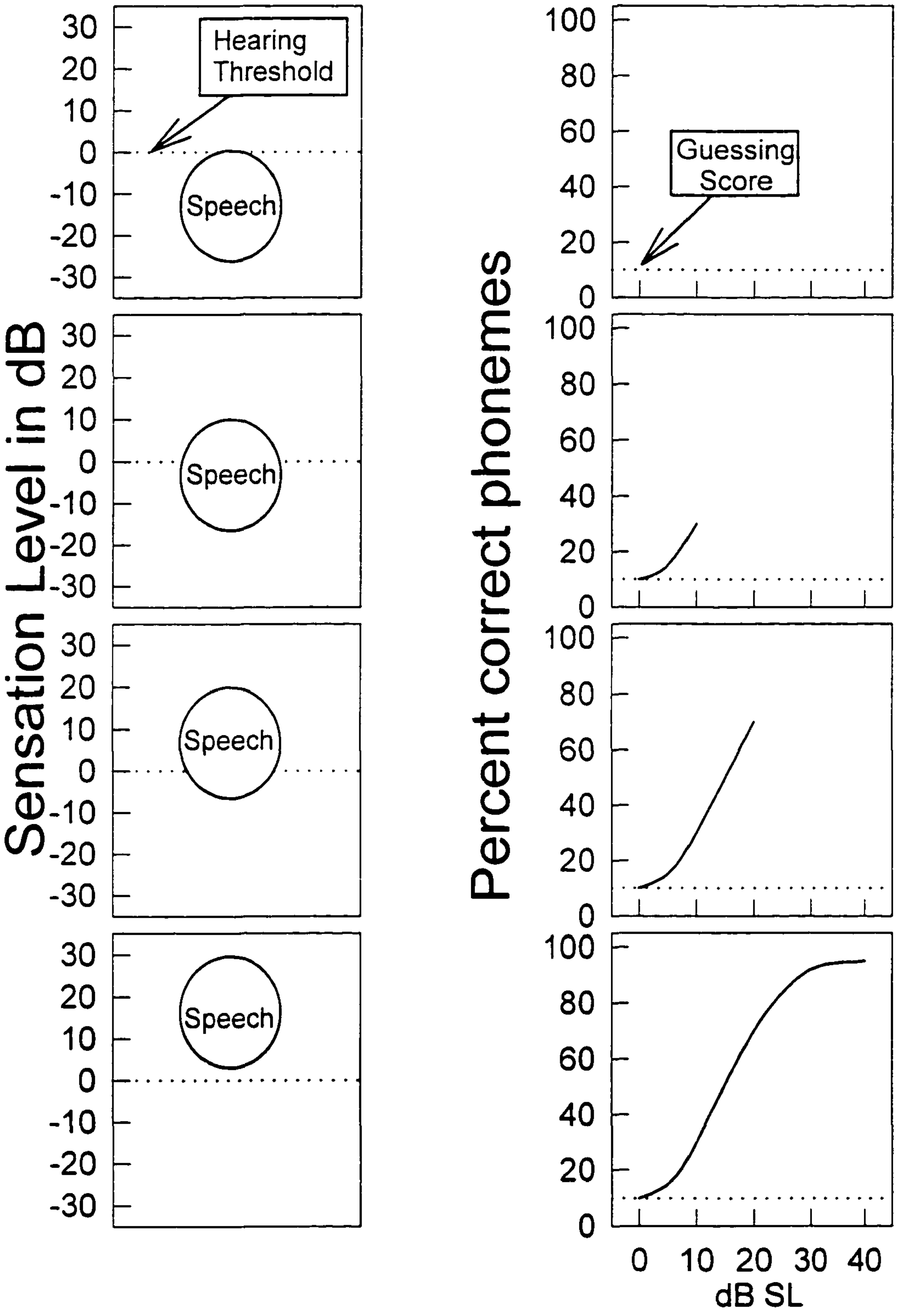

Figure 1.2: The left panels show the audibility of the speech signal as its level is raised. The right panels show their corresponding effect on the PI function.

To quantify the effectiveness of the compression schemes, the perceptual effective compression ratio was calculated and compared against the target compression ratio. The perceptual effective compression ratio was defined here as the ratio of the perceptual dynamic range of the speech signal before compression to that after compression. In the present study, the perceptual effective compression ratio was calculated for different criteria of the perceptual dynamic range of speech. Figure 1.3 illustrates the calculation process of the effective compression ratio. In this example, the perceptual dynamic range of the speech signal (between the 10% to 90% correct recognition points) before processing was about 30 dB. A compression scheme, with a target compression of 3:1, has reduced this range to 15 dB. Therefore, the perceptual effective compression ratio was 2:1.

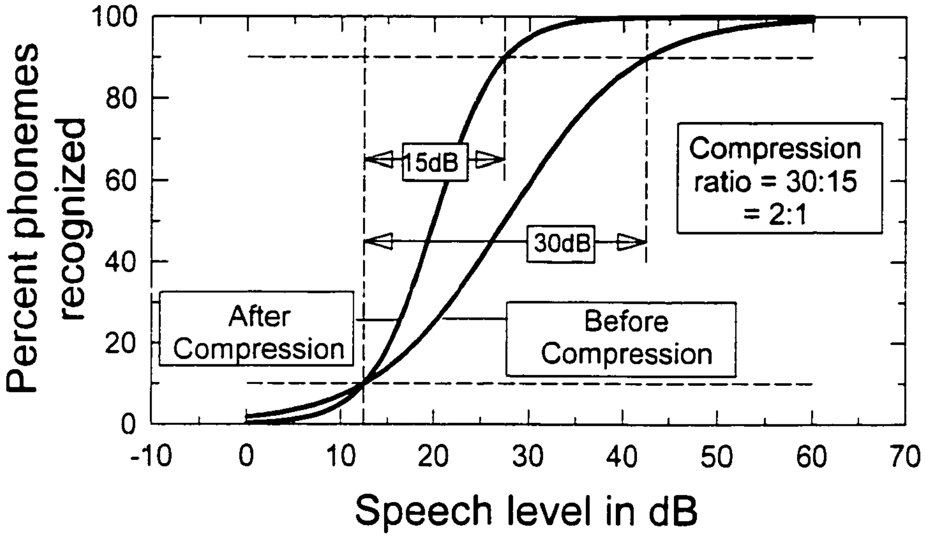

Figure 1.3. Illustrating behavioral estimation of compression ratio, based on the perceptual dynamic range - as measured in this example between the 10% and 90% points of the Performance/Intensity function.

CHAPTER II: REVIEW OF RELATED LITERATURE

2.1. INTRODUCTION:

In this chapter, literature on three areas relevant to compression hearing aids is reviewed. In the first section, the acoustical properties of the speech signal pertinent to compression amplification are discussed. Included here are long-term average spectrum of speech, overall level of speech, and dynamic intensity range of speech. The second section deals with loudness perception including the methods of measurement of loudness, loudness growth in normal-hearing individuals and loudness growth in individuals with sensorineural hearing loss. In the final section, past research in compression amplification is reviewed including the rationale for amplitude compression in hearing aids, the characteristics and classification of compression systems, a review of research findings with WDRC/syllabic compression amplification with emphasis on research on single-band and 2-band systems, and a critique on negative findings with syllabic compression.

2.2. ACOUSTICAL PROPERTIES OF THE SPEECH SIGNAL:

2.2.1. Long-term average spectrum of speech (LTASS):

The long-term average spectrum of speech (LTASS) is a direct measure of the average sound pressure level of speech as a function of frequency. This measurement provides a means of viewing the average frequency-intensity distribution in a continuous speech sample. The measurement of the LTASS is, effectively, made by passing the speech energy through a series of contiguous bandpass filters and integrating the energy at the output of each filter.

Dunn and White (1940) reported group averaged LTASS of 11 talkers (6 male and 5 female) who read a 2.5 minute passage. A total of 600 samples of 125 ms duration each were obtained at the rate of four samples per second from the entire 2.5 minute passage. The width of the analysis frequency bands were 1 octave below 500 Hz, and 1/2 octave above 500 Hz. For each frequency band, the RMS sound pressure level was calculated for each of the 600 samples. The short-interval RMS values were then integrated across all the sampled intervals for each talker to provide the long-interval estimate. The long-interval estimates were averaged across talkers, converted into spectrum level, and plotted on a log-log scale depicting spectrum level as a function of frequency. The result was the average power spectrum for an interval equivalent to 1.25 minutes of speech, in each of 12 frequency bands. The function shows a maximum around 500 Hz. Above this frequency, the spectrum level (level per cycle) rolls off at about 9 dB/octave towards the high frequency region (Olsen, Hawkins, and Van Tasell, 1987). Therefore, the intensity level of speech is much lower in the high-frequency region compared to that in low-frequency region.

A number of factors are known to affect the estimate of the LTASS. These factors can be broadly classified into “method of analysis” factors and “speech signal” factors (Olsen et al., 1987). Method of analysis factors include filter bandwidth, length of the sampled interval, the distance and azimuth of the microphone relative to the talker, speech recording environment, whether or not the analysis included silent intervals between words, pauses, and sentences, and whether the results are reported in terms of spectrum level or overall level in each analysis band. Speech signal factors include sex, age and vocal effort of the talker, and the choice of the speech material (Formby and Monsen, 1982; Stevens, Egan, and Miller, 1947; Dunn and White, 1940; Byrne, 1977; Benson and Hirsh, 1953; Cox and Moore, 1988; Dunn and Farnsworth, 1939; Cornelisse, Gagne, and Seewald, 1991; Niemoeller, McCormick, and Miller, 1974; Tarnoczy, 1956; Pearsons, Bennett, and Fidell, 1977; Brand, Ruder, and Shipp, 1969; Schwartz, 1970).

For hearing aid design and prescription purposes, Byrne (1977) suggested that LTASS should be expressed in units comparable with the resolving capacity of the human ear, namely in critical bands. Since the critical bandwidths of the human ear approximate 1/3 octave bands, 1/3 octave wide filters can also be used for the purpose of hearing aid selection and evaluation. According to ANSI (1969) standards, stable long-term spectra can be obtained with integration times as short as a minute. It is generally agreed that the LTASS for hearing aid selection and evaluation purposes should be obtained for normal conversational speech produced with normal vocal effort. Pearsons et al. (1977) and Cox and Moore (1988) reported LTASS obtained with such comparable parameters for American English speakers. The LTASS of these two studies are in close agreement with each other. The LTASS shows a prominent peak at around 500 Hz and the spectrum rolls-off at about 6 dB/octave above this frequency.

2.2.2. Overall level of speech:

The overall level of speech is obtained by integrating the speech energy over its frequency bandwidth. There are a number of factors that can influence the estimate of the overall level of speech. These include the distance, vocal effort, sex and age and other characteristics of the talker, and recording environment. For practical and hearing aid prescription purposes, the overall RMS level of speech at 1 meter distance from the talker in typical room environment is considered to be about 70 dB SPL (Byrne, 1977; Olsen et al., 1987; Cox and Moore, 1988).

2.2.3. Dynamic range of speech:

The intensity level of the speech signal varies over time. Some of the intensity variations take place slowly while others take place at a much faster rate. The slow intensity variations of the speech signal are generally associated with the variations of the overall level of speech. The overall level of speech can vary over at least a 10-20 dB range due to factors such as vocal effort and talker characteristics and even more when distance is taken into account (e.g., Pearsons et al., 1977; Boothroyd, 1993). In addition to the variation of the overall level of speech, there is also variation of the speech levels due to the characteristics of individual speech sounds. These variations take place over short time intervals and are referred to as short-term dynamic range of speech.

2.2.3.1. Short-term dynamic range of speech:

The short-term dynamic range of speech can be measured both acoustically and perceptually. The acoustical measurement can be further classified into two types. In the first method, the energy of a speech sample is integrated over short time intervals. These short interval levels are then averaged arithmetically to derive the short-term spectral distribution. We shall refer to this method as LTASS method. The other acoustical method is referred to as Phonemic Spectral Distribution. In this method, only the intensities and frequencies of key elements of the speech signal considered important for speech perception are measured. The perceptual measurement of the short-term dynamic range of speech is obtained by measuring percent correct recognition of the speech signal (e.g. percent words recognized or percent phonemes recognized) as a function of the level of the speech signal relative to the threshold of audibility.

1. **LTASS method:** Dunn and White (1940) measured the RMS sound pressure of 600 samples of speech, of duration 125 ms each, in different frequency bands. The 600 short-term levels were then averaged. They also computed the sound levels exceeded in different percentile points (1% to 80%) of the 600 speech samples. The range from lowest percentile point to the highest percentile point was referred to as the dynamic intensity range of speech. According to Dunn and White's data it is about 30 dB. In obtaining the short-term dynamic range of speech

using LTASS method, we need to consider 2 important issues: the integrating time window and the percentile points used to define the range.

Pavlovic (1993), and Cox, Matesich, and Moore (1988) compared the short-term spectral distribution obtained with different integrating times. Pavlovic (1993) used 13 ms, 80 ms, and 200 ms integrating times. Cox et al. (1988) used 20, 40, 60, 80, 100, and 120 ms integrating times. Both these investigators showed that as the integrating time window is reduced, the dynamic range of speech increases. For hearing aid design and prescription purposes, Pavlovic (1993) suggested that the integrating time window should correspond to the temporal integration window of the human ear, i.e., about 100 ms. With respect to choice of percentile points, the dynamic range is defined from the 1st or 10th percentile point to the 90th percentile. These percentile points were arbitrarily defined and their validity has not been studied yet. The Cox et al. (1988) data for 120 ms integration time window show a dynamic range of about 45 dB (1% to 90%). The dynamic range is not constant over the entire frequency range. It is larger (50 dB) in the lower frequency region (250-1000 Hz) and smaller (40) in the high frequency region (1000-6300 Hz). Pavlovic (1993) also reported similar results.

2. **Phonemic spectral distribution of speech:** The LTASS method of defining the spectral distribution of speech has certain drawbacks. As Boothroyd, Erickson, and Medwetsky (1994) have pointed out, LTASS method does not account for inter-talker differences, it ignores the effects

of vocal effort and microphone location on the spectral distribution of speech, it contains no direct information about the relative importance of different spectral regions, and it might underestimate the short-term levels of less frequently occurring speech sounds, specifically the high frequency speech sounds. An alternative approach to the LTASS is the phonemic spectral distribution of speech (Fletcher, 1953; Fant, 1974; Medwetsky and Boothroyd, 1991; Boothroyd et al, 1994). In this approach, the intensities and frequencies of key acoustic features of individual speech sounds are measured.

Fletcher (1953) measured the power of individual speech sounds of CVC words. The relative power of all speech sounds (re. the faintest sound /θ/) were presented in a table. Conversion of this power ratio into dB revealed that the range from the loudest sound (vowel /ɔ/) to the faintest sound (consonant /θ/) was 28 dB (Gerber, 1974).

Boothroyd et al. (1994) measured the short-term spectra (51.2 ms duration, 1/3 octave bands) of 7 speech sounds: /m,u,i,ɑ,s,ʃ,f/. These sounds were selected since they more or less define the extreme frequencies and intensities of speech. It was assumed that the information of all other speech sounds would fall within the boundary defied by these speech sounds. The dB range from the strongest peak (F1 of /u/) to the softest peak (1 peak of /f/) was found to be about 33 dB. This study also revealed substantial inter-subject and gender variability.

The inter-subject variability was larger than inter-gender variability. When individual variability was also considered, the dynamic range of speech increased to about 53 dB.

3. **Performance-Intensity (PI) functions:** The short-term dynamic range of speech can also be measured perceptually by means of PI functions. These functions are also referred to as articulation curves or articulation/intelligibility curves or articulation-gain curves. Typically in this method, CVC words are presented at different levels relative to hearing threshold (absolute or masked) and at each level percent correct recognized is measured. Ideally, the presentation levels are varied to obtain speech recognition scores ranging from a guessing score (about 10% phonemes correct) to a perfect score (about 100%). The intensity range required to obtain scores ranging from a guessing score to a near-perfect score is defined as the "perceptual dynamic range of speech". This range, for a single talker speech produced at a fixed distance and fixed effort, is about 30 dB for CVC words, when measured using normal-hearing individuals (Boothroyd, 1993).

Davis and Silverman (1960) presented data showing percent words recognized as a function of presentation level. Their data clearly showed that, for normal hearing individuals, as the presentation level of CVC words is raised above the threshold of audibility to about 30-40 dB, the percent recognized scores increases and then plateaus. It should be noted that even in normal hearing individuals, the increases in percent recognized with increases in presentation level is not a linear function. The function first grows gradually to about 20% score, then accelerates to about 70-80% scores and then plateaus to about 95-100% scores. The function from the 20% score to about 70% score is, however, more or less linear.

A number of factors have been identified which influence the speech recognition scores. With respect to the measurement of PI function, the following factors are most crucial:

1. **Test Material:** Use of different test material can lead to variations in PI function. PI functions for normal-hearing individuals reported by Davis and Silverman (1960) show that the function slope is steepest for spondees followed by PB W-22 list and PB Hughes recording in that order. Their data showed that the 50% recognition score for spondees is about 8 dB lower than that for PB W-22 lists and about 14 dB lower than that for the PB Hughes recording. Niemeyer (1965) reported that the PI functions for sentences are much steeper than for monosyllabic words. Generally, as the redundancy of the test material increases, the slope of the PI function also increases. Many investigators have suggested the use of CVC words for tests of speech recognition as they contain the least amount of redundancy. Boothroyd (1968) suggested scoring of CVC words at phoneme level to further reduce the effects of redundancy.

Olsen and Matkin (1979) compared PI functions reported by different investigators for different speech material. Even for CVC words, they noted that while the functions reported by the different studies are similar to each other, none of them directly superimpose upon each other. In general, the slopes in the linear portion of the functions varied from about 3.8%/dB to about 5.6%/dB.

2. Word Familiarity: The effect of difficulty and familiarity of words on speech recognition has been well documented in the literature. Penrod (1994) noted that speech recognition tests incorporating words that are not within the vocabulary of the individual who is being tested can lead to spuriously low scores, leading to unwanted testing, misdiagnosis and mismanagement. Owens (1961), and Epstein, Giolas, and Owens (1968) have shown that as the familiarity of words increase, even to a slight degree, speech recognition scores for phonetically balanced CVC words increase. At the sentence level, Duffy and Giolas (1974) and Giolas, Cooker, and Duffy (1970) have shown that percent recognized of key words is higher for high predictability items than low predictability items.

3. List Size: Half lists of 50-item lists have been recommended in an effort to reduce the test time and avoid patient fatigue. Some studies reported good agreement when 2 half-list scores were compared and when half-list scores were compared with full-list scores (Elpern, 1961; Resnick, 1962). In contrast, Jirsa, Hodgson, and Goetzinger (1975) and Schwartz, Bess, and Larson (1977) recommended against the use of half-lists. It has been pointed out in the literature that when the sample size is reduced, the variability and reliability of obtained scores are affected. Using the binomial model, several investigators have pointed out that as the test size is reduced, the variability in scores increases, and the farther the scores are from 0% or 100%, the less reliable the scores are (Boothroyd, 1968; Gelfand, 1998; Hagerman, 1976; Raffin and Schafer, 1980; Raffin and Thornton, 1980; Thornton and Raffin, 1978). Boothroyd (1968) suggested scoring at the phoneme level to increase the number of test items by a factor of 3 for CVC words. Gelfand (1998) suggested that speech recognition tests should contain approximately 450 items in order to attain adequate reliability of a score.

2.2.4. Relationship between LTASS and short-term dynamic range of speech:

Dunn and White (1940), Cox et al. (1988), and Pavlovic (1993) studied the relationship between the LTASS and the short-term dynamic range of speech. Generally, the peak levels of speech (1% levels) lie about 12 dB above the LTASS. The 90% levels lie about 18 dB below the LTASS. Similar results were also reported by Boothroyd et al. (1994) when they plotted the phonemic spectral distribution of speech against the LTASS reported by Cox and Moore (1988).

To summarize thus far, speech is a dynamic signal. For hearing aid design and selection purposes, we need accurate information about 3 fundamental aspects of the speech signal, namely the shape of the average spectrum, the overall level of speech and the dynamic range of speech. There are a number of variables that affect the estimate of the LTASS. It is generally agreed that LTASS should be obtained for conversational speech samples produced with normal vocal effort of at least 1 minute duration using 1/3 octave band filters. When spectra obtained in this manner are plotted on a log-log scale depicting overall level in each 1/3 octave band as function of frequency, the familiar spectrum shows a maximum at about 500 Hz. Above this frequency the spectrum rolls-off at roughly 6 dB/octave. The overall level of speech is obtained by integrating the speech energy over its bandwidth as well as a finite time period. The overall RMS level of conversational speech spoken with normal vocal effort, at a distance of 1 meter form the speaker in typical room environments is considered to be about 70 dB SPL. There are two kinds of intensity variations of the speech signal: slow variations and fast variations. The slow intensity variations of the speech are influenced by variations of vocal effort, distance, and talker characteristics. They cause the overall level to vary over a 10-20 dB range. The fast intensity variations of the speech are influenced by the properties of the individual speech sounds. These are referred to as short-term dynamic range of speech. The short-term dynamic range of speech is about 30-40 dB. When this range is added to the intensity range of the overall level, the dynamic range of speech is about 50-60 dB or even more.

2.3. LOUDNESS PERCEPTION:

2.3.1. Methods of measurement of loudness:

The methods of measurement of loudness can be classified into direct and indirect methods. Magnitude estimation, magnitude production, and categorical scaling are direct methods of measuring loudness functions. Loudness balancing/matching and cross-modality matching are the indirect methods of measuring loudness.

1) Magnitude estimation: In this method, a stimulus is presented at various levels and the subject is required to assign to each stimulus a positive number corresponding to the subjective loudness of that stimulus. In the earlier work by Stevens (1955), a standard stimulus was presented and an arbitrary number was assigned to its loudness, e.g. 1 to 100. Then a comparison stimulus was presented and the subject was instructed to assign a number to the loudness of the comparison stimulus. This is referred to as free magnitude estimation (Launer, 1995).

A variation of magnitude estimation is called absolute magnitude estimation. In this method there is no comparison stimulus. The subjects are instructed to merely assign any positive integer to the perceived loudness (Hellman and Meiselman, 1993).

2) Magnitude production: In this method, the subjects are presented with a standard stimulus and asked to adjust the level of a test stimulus until its loudness matches a specified loudness.

Free magnitude estimation and magnitude production methods rely on the ratio scale. Stevens (1957) believed that subjects estimate the loudness of a signal using a ratio scale. That is, a given sound may be judged to be twice as loud as another or half a loud as another. In contrast, Pascoe (1978) proposed a categorical scale for the measurement of loudness.

3) Categorical scaling: This method assumes that the listener can partial the dynamic range into verbal categories such as soft, comfortable, or loud. In this method, stimuli are presented at different levels and the subject is instructed to assign a verbal category from a predefined scale (Allen, Hall, and Jeng, 1990).

4) Loudness balancing/matching: In this method, the listener compares the loudness of two stimuli and adjusts one of them in level to produce equal loudness. Generally, the intensity of one of the stimuli is fixed in level (reference) while the other is varied in level (target) until both stimuli are judged to have the same loudness. This is repeated for several levels of the reference sound.

When the reference and the target stimuli are presented alternatively to opposite ears, this method is called *Alternate Binaural Loudness Balance* (ABLB). When the reference and the target stimuli are presented to the same ear, the method is called *Monaural Loudness Balance* (MLB). ABLB is used as a diagnostic test for loudness recruitment in unilateral hearing loss. MLB is used for construction of the equal loudness contours. It is also used as a diagnostic test for loudness recruitment in bilateral hearing losses where there is normal hearing for at least one frequency in the test ear.

2.3.2. Loudness growth in individuals with normal hearing:

Loudness function refers to the relationship between the intensity of the signal and its perceived loudness. A number of variables related to both the stimulus and the listener are found to affect the nature of the loudness function (for a detailed review of these factors, see Scharf, 1978).

The loudness function for a 1000 Hz tone is often referred to as the "standard" loudness function. Stevens defined the loudness of a 1000 Hz pure tone at 40 dB SPL as one sone. He also defined the 40 dB SPL as 40 dB phon level. The loudness level in phons is the SPL of an equally loud 1000 Hz tone. Thus, the loudness level in phons of a 1000 Hz tone is the same as its SPL in dB. The loudness function is based on direct psychophysical procedures such as magnitude estimation (Stevens 1955, 1957). The loudness function shows loudness in sones as a function of loudness level in phons.

The equation for the standard loudness function is a simple power law given by

$L = KP^{0.6}$; where 'L' is the loudness, 'P' is the sound pressure and 'K' is a constant, or

$L= KI^{0.3}$; where 'I' is the intensity (Stevens, 1957).

According to the power law, a 10 dB increase in SPL of the signal would result in doubling of loudness. Near absolute threshold (0-30 phons), the loudness function grows faster. As a result, the function is curved near threshold and can be approximated as

$L = K (P-Po)^{0.6}$; where Po is the effective threshold (Scharf, 1978).

In normal hearing individuals, the loudness function is steepest for the low frequencies, less steep for the high frequencies, and even less steep for the mid frequencies.

2.3.3. Loudness growth in individuals with sensorineural hearing loss:

Fowler (1936) was among the first to describe loudness intensity relations in subjects with unilateral sensorineural hearing loss. He showed that in these individuals loudness grows rapidly in the impaired ear. He coined the term "Recruitment". He noted that recruitment is a characteristic of sensorineural hearing loss and that it is absent in individuals with pure conductive hearing losses. Historically, researchers relied on loudness balance tests to investigate recruitment in individuals with sensorineural hearing loss.

1) Loudness balance: Harris (1953) critically reviewed the literature on loudness recruitment. He classified recruitment into 5 types: No recruitment, complete recruitment, partial recruitment, over recruitment, decruitment. This classification is based on the shape of the ABLB curves. In "no recruitment", seen in individuals with unilateral conductive hearing loss and in some individuals with retrocochlear pathology, the shape of the loudness balance curve in the impaired ear parallels that of the normal ear. The distance between the normal and impaired curves determines the degree of hearing loss in the impaired ear. “Complete recruitment” is seen in individuals with unilateral sensorineural hearing loss where the loudness in the impaired ear grows rapidly as the intensity is increased from just above the threshold to about 30 dB SL. At some supra-threshold level, the loudness growth in the impaired ear catches up with that of the normal ear. In individuals with mixed hearing loss, loudness rapidly grows as the level is increased above threshold and at some level above threshold the rate of growth of loudness becomes linear. However, the linear part of the impaired ear's curve never catches up with that of the normal ear. It merely parallels the normal ear's curve. This type of recruitment is called "partial recruitment". In some individuals with sensorineural hearing loss, "over recruitment" is seen. Here, the loudness grows faster in the impaired ear, catches up with that of normal ear at some supra threshold level and continues to grow faster than that in the normal ear. Decruitment is sometimes noted in individuals with retrocochlear pathology, in which an increase in stimulus level causes a reduction in loudness. This presumably happens due to decay of neural activity. Jerger (1962) also proposed similar classification.

Hallpike & Hood (1959) obtained loudness balances (ABLB) from a group of 200 individuals with unilateral sensorineural hearing loss (Meniere's disease). They noted that all of the individuals exhibited complete recruitment, that is, at some supra-threshold level the perceived loudness in the impaired ear equaled that in the normal ear. Further they noted that two-thirds of the subjects exhibited some degree of over recruitment. In order to quantify recruitment they measured the recruitment angle defined as the angle between the recruitment curve and the horizontal axis of the graph. The recruitment angles for 76 individuals were measured for 500, 1000, 2000, and 4000 Hz as a function of degree of hearing loss. The data for each frequency were fitted with regression lines with confidence limits of +/- 2 SD. On average the recruitment angle increased by approximately 1 degree for each 2 dB increment of hearing loss. That is, as the degree of hearing loss increased the rate of loudness growth also increased.

Hallpike (1967) obtained loudness balances (ABLB) in a group of 50 individuals with unilateral sensorineural hearing loss. Among the 50 subjects, 30 had Meniere's disease and 20 had 8th nerve tumors. They noted that all 30 subjects with Meniere's disease exhibited complete recruitment. In 14 of the 20 subjects with 8th nerve tumor, recruitment was absent; in the remaining 6, partial recruitment was found. Thus, they noted that the presence of recruitment is a strong indicator of cochlear pathology. However, they noted that even individuals with retrocochlear pathology (about 10%) may exhibit some degree of recruitment.

Miskolczy-Fodor (1960) obtained binaural loudness balances in 100 subjects with unilateral sensorineural hearing loss for 1000 Hz and 4000 Hz test frequencies. The subjects with hearing loss were divided into 4 categories depending upon the degree of hearing loss at the test frequency: 40, 50, 60 and 80 dB HL. All subjects exhibited complete recruitment. Further, the slope of the loudness balance was directly proportional to the degree of hearing loss, that is, as the degree of hearing loss increased, the mean slopes also increased.

2) Loudness functions: Knight & Margolis (1984) obtained loudness-intensity functions using the magnitude estimation method in a group of 5 young adults (20-42 years) who had asymmetrical sensorineural hearing loss. Absolute magnitude estimates were obtained in the 250-6000 Hz frequency range (at octave and some inter-octaves). At each frequency, magnitude estimates were obtained at 5 stimulus levels with 5 replications. The results of the impaired group were compared to that of a group of 10 normal hearing subjects. All of the impaired subjects showed rapid loudness growths as intensity was increased. The slope of the loudness functions at 30 dB HL was calculated by fitting each function with a third-order polynomial. The slope of the loudness functions in the sensorineural hearing loss group exceeded the 95% confidence limits for that of the normal hearing group. No attempt was made to determine the effect of degree of hearing loss on the slope of the loudness functions.

Hellman & Meiselman (1990) investigated the effect of degree of hearing loss on the slope of the loudness function in a group of 100 individuals with sensorineural hearing loss. The age range of the subjects was 19-80 years. Seventy eight of the subjects studied had bilateral cochlear hearing loss due to noise exposure. These 78 subjects were divided into 4 groups based on their thresholds: 45, 55, 65, 75 dB HL. Absolute magnitude estimates of loudness were obtained for an average of about 35 dB SL above thresholds. Frequency was not an independent variable in this study. That is, different subjects were tested at different frequencies in 500 to 4000 Hz range. The only criterion was that they should have a hearing threshold of not less than 40 dB HL, presumably due to the same etiology. The slopes of the hearing impaired groups were compared with that of the normally hearing group of 51 subjects. As expected the slope of the loudness functions increased with increasing degree of hearing loss. Hellman and Meiselman (1993) also reported similar results for loudness functions obtained using magnitude production and cross modality matching methods.

Launer (1995) also investigated the effect of degree of hearing loss on the slopes of the loudness functions obtained using categorical scaling method in a group of 67 hearing-impaired subjects. Like Hellman and Meiselman (1990, 1993), this study also showed increases in the slope of the loudness function with increases in the degree of hearing loss.

To summarize thus far, loudness is a monotonic function of stimulus intensity in normal hearing individuals. It can be measured using different methods: magnitude estimation, magnitude production, cross-modality matching, loudness balances, and categorical scaling. Fowler (1936) was among the first to demonstrate that subjects with sensorineural hearing loss tend to exhibit more rapid loudness growth with increasing signal level than normally hearing subjects. Using loudness balance techniques, different investigators have shown that recruitment can be classified into different categories based on the shape of the loudness functions. These are no recruitment, complete recruitment, partial recruitment, over recruitment and decruitment. Complete and over recruitment are hallmarks of sensorineural hearing loss. The slope of the loudness functions in sensorineural hearing loss is steep in comparison with normal hearing. The slope tends to increase with increasing degree of hearing loss.

2.4. AMPLITUDE COMPRESSION IN HEARING AIDS:

2.4.1. Rationale of amplitude compression:

Amplitude compression has been incorporated in hearing aids for purposes including avoidance of discomfort and distortion, reduction of the short-term dynamic intensity range of speech, reduction of long-term dynamic intensity range of speech, loudness normalization, and noise reduction (Dillon, 1996). These rationales will be briefly described.

Amplitude compression is often used in hearing aids to limit the maximum output that it can generate. Individuals with sensorineural hearing loss often exhibit elevated thresholds of audibility but normal or lower than normal thresholds of auditory discomfort. With linear amplification, there is a high possibility of the hearing aid's output exceeding the threshold of discomfort. When an ear is stimulated with levels higher than thresholds of auditory discomfort, there is a potential for further noise-induced hearing loss (Macrae, 1991). Also, hearing aid users are known to reject a hearing aid if its output often exceeds threshold of discomfort (Wallenfels, 1967). For these reasons, therefore, the maximum output a hearing aid can deliver is often limited at levels below the threshold of discomfort.

A simple form of output limiting is peak clipping. Unfortunately, peak clipping generates harmonic and inter-harmonic distortions. These distortions can affect speech intelligibility and sound quality (Craig and Van Tassell, 1994). Compression limiting is often a preferred choice of output limiting since it is essentially free from harmonic and inter-harmonic distortions.

In individuals with sensorineural hearing loss the dynamic range of hearing is reduced. In severe and profound hearing loss individuals, the dynamic range of hearing is often much less than the short-term dynamic range of speech. With linear amplification, these individuals are restricted to listening at low sensation levels of speech and, therefore, miss low intensity speech elements which may be important for speech understanding. With amplitude compression, more low intensity elements of the speech signal can be fit into the reduced dynamic range of hearing and possibly improve speech understanding.

The speech recognition scores are dependent on the overall level of speech. As the overall level of the speech signal is varied, a corresponding change in speech recognition score is noted. This effect is noted in both normal hearing and hearing-impaired individuals. With amplitude compression, variations of the overall level of speech can be reduced so as to maintain optimal speech perception.

Amplitude compression in hearing aids has been advocated to compensate for the effects of loudness recruitment. That is, amplitude compression is used as a means for normalization of the rather rapid growth of loudness in individuals with sensorineural hearing loss. For loudness normalization, low input levels are amplified more than high intensity levels, and for levels at which loudness in the impaired ear corresponds to that of the normal ear, no amplification is provided. Since recruitment is more pronounced at high frequencies, amplitude compression for loudness normalization is often applied only in the high-frequency band.

Individuals with sensorineural hearing loss experience increased speech understanding difficulty in background noise compared to normal hearing individuals (Olsen and Tillman, 1968). They require greater signal-to-noise ratios for optimal speech understanding. The goal of amplitude compression for noise reduction is to reduce gain when increases in input are due to background noise. In general, noise is weighted more toward low frequencies. Therefore, amplitude compression for noise reduction is applied only in the low-frequency band. It is assumed that increase in input levels in the low-frequency band is due to background noise and therefore its gain is reduced. Thus, compression is applied only for high intensity input levels in the low-frequency band.

2.4.2. Characteristics and classification of compression hearing aids:

Most typical compression systems operate by means of a feedback loop. In such systems, a level detector estimates the signal level at some point in the amplification process and feeds back the information to the compression amplifier to adjust its gain for subsequent inputs. Because of the feedback loop, there will always be a time lag between the level detection and compression activation. We shall refer to such systems as look-back compression systems. It is also possible to design a compression system that operates on a feed-forward loop. In such systems, a signal delay pathway is introduced so that a level detector that estimates the signal level can be used to adjust the gain of the signal per se. We shall refer to such compression systems as look-forward compression systems.

Characteristics: A compression system is defined by its operating characteristics. These characteristics can be broadly classified into two types: static characteristics and dynamic characteristics.

A. **Static characteristics:** The static characteristics of a compression system can be described using input/output (I/O) function. While several terms are used to describe the I/O curve, the following 2 are most important for classification purposes:

1. **The compression threshold:** It is the minimum input level required to activate the compression function of the system and is defined as that input level at which the gain is 2 dB less than that in the linear region. The region below the compression threshold is called the linear region and that above the compression threshold is called the compression region.
2. **Compression ratio:** It is the inverse slope of the I/O curve. That is, it is the ratio of small change in input level to the corresponding change in the output level.

B. Dynamic characteristics: The attack time and the release time typically describe the dynamic characteristics of a compression system. They roughly describe the output envelope after a given change in the input envelope.

1. **Attack time:** It is the time required for the output of a compressor to come within 2 dB of the level specified by the I/O curve after the input level increases by at least 25 dB to a level above the compression threshold.
2. **Release time:** It is the time required for the output of a compressor to come within 2 dB of the level specified by the I/O curve after the input level decreases by at least 25 dB to a level below the compression threshold.

Classification: Compression hearing aids can be classified in a number of ways. They include classifications based on whether or not a feedback loop is used, if a feedback loop is used - its position relative to the volume control, the intensity range over which compression is applied, the number of frequency bands in which the compression is applied, the speed with which the compression is applied, and the magnitude of the compression ratio.

It was mentioned earlier that compression systems could be classified into look-backward or look-forward compression on the basis of utilization of the feedback loop. Look-backward compression systems can be further classified into input compression and output compression based on the relative position of the feedback loop and the volume control. When the feedback loop is located before the volume control, it is known as input compression. On the other hand, when the feedback loop is located after the volume control it is known as output compression. For input compression, changes in volume control do not alter the compression threshold but affect the output at which compression is activated. For output compression, changes in volume control do affect the compression threshold but not the output level at which compression is activated. For these reasons, input compression is recommended for syllabic compression and output compression is recommended for compression limiting.

Walker and Dillon (1982) provided a classification based on the intensity range of compression, the speed of compression and the magnitude of compression. Based on these 3 operating characteristics, they essentially came up with 3 forms of compression, namely, compression limiting, wide dynamic range syllabic compression, and slow-acting automatic gain control (AGC). Compression limiting is characterized by high compression threshold, high compression ratio, and short attack and release times. This type of compression is used to limit the maximum output a hearing aid can generate. Wide dynamic range syllabic compression is characterized by low compression threshold and short attack and release times. While compression thresholds are usually low, high compression ratios can be used, especially for individuals with severe and profound hearing losses. This type of compression is used to reduce the short-term dynamic range of speech. In slow-acting AGC, the compression threshold is low, and the attack and release times are long. Compression ratio can be low or high. This form of compression is intended to reduce the long-term dynamic intensity range of speech.

Dillion (1996) classified the currently available compression systems based on the intensity range of compression, the frequency bands of compression, and the speed of compression. This classification system is especially useful in characterizing the different compression schemes as they relate to the underlying theoretical rationale. Inspection of this classification system, again points to essentially three forms of compression namely compression limiting, AGC, and syllabic compression. Of course, the nature of syllabic compression used varied significantly in their classification system. Complete details will be found in the original article.

The present study focused on wide dynamic range syllabic compression. Note that, in this paper, the terms WDRC and syllabic compression have been used interchangeably to refer to the reduction of the short-term dynamic range of speech to improve speech intelligibility. Past research in this area will now be reviewed.

2.4.3. Review of past literature on WDRC/syllabic compression:

Braida, Durlach, Lippman, Hicks, Rabinowitz, and Reed (1979) provided a detailed critical review of literature on syllabic compression. They noted that prior to the 1970's most of the studies concentrated on single-band compression. In general, the studies failed to show clear-cut advantages of single-band syllabic compression. Braida et al. opined that "...the benefits of single-band syllabic compression are likely to be quite limited, however, because the compression curve cannot be varied as a function of frequency and thus compressor action cannot reflect variations in hearing loss with frequency, or changes in the spectral characteristics of the input signal". Further, many of these studies were confounded by limitations such as use of very high compression ratios, inappropriate attack/release times, system-induced noise and distortions, and poor design. To overcome the limitations imposed by single-band compression systems, many investigators have proposed using multi-band compression systems in which different amounts of compression can be applied in different frequency bands to tailor to individual subject's needs.

Villchur (1973) evaluated a 2-band syllabic compression system in a group of 6 subjects with moderate to severe sensorineural hearing losses. The crossover frequency between the 2 bands varied from 1.3-2.5 kHz, across the subjects. Compression ratio and equalization were selected in an attempt to restore normal loudness relationships for speech. Subjects were allowed to fine tune the compression ratio based on preferred judgments. On average the compression ratio was 2.1:1 for the low-frequency band and 2.8:1 for the high-frequency band. This system was compared to a linear amplification system that had flat frequency response above 500 Hz. Speech recognition scores for nonsense CVC words, presented within a sentence, were obtained for both quiet and in a 10 dB signal-to-noise ratio. Stimuli were presented at MCL and at a 10 and 20 dB reduced input level to the compression system.

In quiet, the recognition score with the 2-band compression was better than linear amplification especially for the reduced input levels. While this study showed a significant advantage for 2-band compression over linear amplification, especially for low-intensity levels, it was criticized for its choice of the control condition. That is, by comparing the 2-band system against an inferior linear amplification system that had inadequate high-frequency emphasis, this study might have artificially inflated the benefits of 2-band compression system (Braida et al., 1979).

Yanick (1976) compared desktop single-band compression, 2-band compression, and linear amplification. Two groups of subjects were included in this study having flat and sloping hearing losses. For the single-band compression, the average compression ratio was 1.8:1. For the 2-band compression, the average crossover frequency between the 2-bands was 1.5 kHz. For the flat hearing loss group, the average compression ratio was 2.4:1 for the low-frequency band and 2.8:1 for the high-frequency band. For the sloping hearing loss group, the average compression ratio was 1.5:1 for the low-frequency band and 2.5:1 for the high-frequency band. Unfortunately, this study also used a flat frequency response linear amplification control condition. Sentence level stimuli were presented at a subject's preferred level for maximum intelligibility. The stimuli were presented in a background of cafeteria noise at a signal-to-noise ratio of 0 and 6 dB. Comparison of the signal-band and 2-band compression systems revealed higher recognition scores with 2-band compression system than single-band compression system. Further as the S/N ratio was decreased, a higher reduction of recognition scores occurred for single-band than 2-band compression system. Therefore, these results suggest superiority of the 2-band compression scheme over single-band compression scheme, especially under background noise. Unfortunately, meaningful comparison between the compression systems and the linear amplification cannot be made because of poor choice of the linear amplification condition.

Barfod (1976) compared a 4-band linear amplification system with 1-, 2-, and 3-band syllabic compression systems. Five subjects with bilateral steeply-sloping hearing loss were included. All subjects had normal hearing at 500 Hz and below. The linear amplification system was chosen to provide gain in accordance to the hearing loss. The 3 compression systems were chosen to restore normal loudness contours. The 3 compression systems contained 4 frequency bands. For the lowest frequency band, linear amplification was provided. For the other 3 bands, its bandwidth, cut-off frequencies, compression ratio, and the relative gain among the bands were chosen to restore normal equal-loudness contours.

Test material consisted of nonsense CVC words, which were presented in quiet as well as in noise (S/N ratios of -5, 0, 5, 10, and 15). The input level of the speech was fixed at 65 dB SPL. The average scores (averaged across the quiet and the different S/N ratio listening conditions) for the linear amplification condition were superior to the 1-, and 2-band compression conditions and equivalent to the 3-band compression system. Thus, this study failed to show the superiority of any of the 3 compression conditions over the carefully chosen linear amplification condition. However, this study showed that the 3-band compression condition was better than 1-band and 2-band compression systems. It should be noted that the presentation level in this study was fixed at 65 dB SPL. Therefore, the 4 conditions were essentially evaluated at one equal-loudness contour. There are reasons to believe that the benefit of compression, which in theory would normalize more than one equal loudness contour, would be evident by comparing performance at different input levels to the systems.

Abramovitz (1980) compared a 2-band compression system with that of a single-band compression system, a flat frequency response system, and a loudness discomfort level (LDL) based frequency response linear system. The flat frequency response system had uniform gain above 250 Hz. For LDL based response, the frequency response was shaped such that the 1/3 octave band peak levels of the speech lay 5 dB below the subjects' LDL. The non-compression systems were evaluated with and without peak clipping (sharp clipping 10 dB below the speech peaks). The compression ratio for the single-band system was 3:1. For the 2-band compression system, two forms of compression ratios were used. In one condition, the same 3:1 compression ratio was used in both bands. In the other condition, the low-frequency band had a compression ratio of about 1.5:1 and the high-frequency band had a compression ratio of about 3:1. A modified version of the Nonsense Syllable Test (NST) was administered to 4 subjects with moderate-to-severe sensorineural hearing loss. With the exception of the 2-band compression system, all other systems were evaluated both in quiet as well as at a S/N ratio of 10 dB (cafeteria noise).

All systems yielded better results in quiet than in noise. Comparison of the 2 non-compression conditions revealed that the LDL frequency response consistently yielded better performance than flat frequency response in quiet, in noise, without peak clipping, and with peak clipping. Comparison of compression conditions to the LDL frequency response condition revealed that the addition of either single-band or 2-band compression to the LDL frequency response did not improve the speech recognition scores, both in quiet and in noise. However, both single and 2-band compression systems yielded better scores than the flat frequency response linear system. Further, this study revealed no significant advantage of 2-band compression over single-band compression. Thus, this study showed that both single-band and 2-band compression systems are superior to a flat frequency response system but inferior to a LDL-based linear frequency response system. The finding of no benefit of compression over carefully chosen linear amplification in this study could be explained, in part, on the basis of the presentation level chosen in this study. Given that the subjects in this study had only moderate and moderately-severe hearing loss and since the speech material consisted of CVC syllables which may have approximately 30 dB range, the presentation level for LDL frequency response might have resulted in full audibility of the signal and therefore might have maximized performance by itself. Under these circumstances, therefore, the addition of compression to already fully audible speech signal is not likely to show any further benefit.

It is worthwhile to mention that in this study, Abramovitz noted a large inter-subject variability in the data. For some subjects, syllabic compression yielded higher scores than linear amplification. This is a consistent finding of all of the other studies to be reported here. This finding suggests that we need to carefully evaluate compression systems on an individual-to-individual basis before opting for a linear frequency response hearing aid system.

During the early- to mid-1980's investigators evaluated syllabic compression systems in which compression was applied in as many as 16 channels. The impetus to such studies was perhaps derived from the fact that the hearing loss and the dynamic range of hearing vary as a function of frequency and that by using a number of compression bands a closer match between the dynamic range of speech and the dynamic range of hearing could be made. During this time, there was also some evidence to support the notion that speech recognition ability improves as the number of bands of compression is increased (Barfod, 1976).

Lippman, Braida, and Durlach (1981) compared 2 forms of a 16-band syllabic compression system with 4 forms of linear amplification. The four linear amplification conditions were: 1) A flat frequency response (L1), 2) high-frequency emphasis which placed as much of the speech signal as possible between the threshold of audibility and the threshold of discomfort (peak levels at the highest comfortable level)(L2), 3) high-frequency emphasis that restored normal loudness to the 10% peak cumulative amplitude distribution (L3), and 4) mirrored audiogram (L4). The two compression conditions were: 1) less dynamic range compression and high-frequency emphasis than required for normal loudness contours (C1), and 2) compression to restore normal loudness contours for pure tones (C2). On average, the attack time of the compressor was about 1.5 ms and the release time was about 20 ms. The compression ratio for the C1 compression system ranged from 1.0:1 below 500 Hz to about 3.0:1 at the 2000 Hz. A total of five subjects participated in this study. Two of the subjects had flat frequency hearing loss. Inspection of their audiograms revealed a dynamic range of hearing of approximately 30 dB or more. Three of the subjects had sloping hearing losses. The dynamic range of hearing of this group was 60 dB or greater at the low frequencies, gradually decreasing to about 10 dB at high frequencies.

Test material consisted of nonsense CVC words and sentences. Testing was conducted in quiet and in a background of cafeteria noise (S/N ratio was 10 dB). All scores in quiet were superior to all scores in noise, for all conditions. Subjects obtained significantly higher scores with linear amplification conditions L2, L3, and L4 than L1. That is, high-frequency emphasis conditions yielded better scores than a flat frequency response. In general, this study failed to show the benefit of compression over high-frequency linear amplification. This was especially true for the CVC stimuli. For the sentence stimuli, however, the 2 subjects with flat hearing loss obtained higher scores with the C1 compression system than the high-frequency linear amplification system. The authors argued that the benefit of compression would be apparent for a wide dynamic range signal. As stated earlier, the dynamic range of hearing of the 2 subjects with flat hearing loss was approximately 30 dB. This dynamic range should be sufficient for optimal perception of CVC words, which have a narrower dynamic intensity range than sentences.

Comparison of the 2 compression systems revealed that the performance with C1 system was slightly superior over C2. That is, the compression scheme that used lower compression ratios and high-frequency emphasis yielded better scores than the system that restored normal loudness contours. This finding suggest that too much compression could be detrimental to speech recognition. This study also found considerable inter-subject variability in performance with the different systems.

In another experiment of this study, the effects of variations of input level on speech recognition were evaluated. The C1 compression system was compared with the best linear system for a given subject as noted in the above experiment. The input level to compression and linear systems were set at 0, 8, 16, and 24 dB below the levels corresponding to each subject's most comfortable level. Percent phonemes recognized in nonsense CVC words were obtained for 4 subjects. As expected, the performance with linear and compression systems were equivalent at the most comfortable level. As the input level decreased, a greater decrease in scores was noted for the linear condition than the compression conditions. This finding reveals that benefit of compression becomes evident when the input levels vary, which always happens in reality.

Similar results of no benefit of multi-band compression over carefully chosen high-frequency emphasis linear amplification, when compression was evaluated at levels biased toward high-frequency emphasis conditions, were also reported by Nabelek (1983) for a 3-band syllabic compression; Walker, Byrne, and Dillon (1984) for a 6-band syllabic compression; and De Gennaro, Braida, and Durlach, (1986) for 16-band compression. On the other hand, in a series of experiments by Moore and his group, 2-band compression was shown to be substantially superior to linear amplification and single-band compression amplification. This led Plomp (1988) to suggest that too many compression bands as well as small attack/release times degrade the spectro-intensity characteristics of the speech signal and thus do not benefit hearing aid users. The studies with 2-band compression will first be reviewed before discussing the issues raised to account for the negative findings with syllabic compression.

Laurence, Moore, and Glassberg (1983) evaluated a high-fidelity 2-band compression hearing aid in both quiet and in noise. The system contained an input slow-acting AGC followed by 2-band syllabic compression. The attack/release times of the AGC was 2/500 ms. The compression threshold of the AGC was set at 70 dB SPL with a range of 20 dB. This AGC ensured that for levels over 70 dB, the overall output level was constant. The output of the AGC was split into 2 bands, with a crossover frequency of 1500 Hz. Each band had one syllabic compressor. The attack/release times were 2/50 ms for the low-frequency band and 2/10 ms for high-frequency band. The outputs of the two bands were mixed and fed to a power amplifier and receiver. When the compression circuits were shut-off, the system acted as a high-fidelity linear hearing aid. The compression system and the high-fidelity linear amplification system were fitted binaurally. The performance with these 2 systems was compared to the subject's own hearing aids or no aids. The high-fidelity linear amplification system was adjusted to provide comfortable loudness listening level for a 70 dB SLP speech input level, in both low- and high-frequency bands. The 2-band compression system was adjusted to provide comfortable loudness listening level for a 70 dB speech input and audibility for a 50 dB speech input, in both the low-frequency and the high-frequency band. A total of 8 subjects with moderate-to-severe sensorineural hearing loss participated in this study. Each subject was tested monaurally and binaurally, in sound field. Sentences were used as stimuli for testing in both quiet and in noise, and were presented from a speaker directly in front of the subject. In quiet, the stimuli were presented at 85, 70, and 55 dB SPL (peak speech level). The dependent variable was percent correct key words recognized.

In quiet, performance with the 2-band compression was superior to that of the high-fidelity hearing aid which in turn was superior to that of the subjects' own hearing aid. At the highest presentation level of 85 dB SPL, there was no significant difference in scores between the 4 systems. When the input speech level was decreased from 85 to 55 dB SPL, speech intelligibility with the 2-band compression system remained essentially unchanged whereas with the other two hearing aids speech intelligibility decreased.

To test speech intelligibility in noise, speech-shaped noise was presented at a level of 65 dB SPL. The speech levels required to obtain 50% intelligibility (SRT) were then measured. Testing was carried out in a sound field. The noise was presented in two ways: in the first way, both the noise and the speech was presented from a single speaker positioned directly in front of the subject (coincidental condition); in the second way, the noise was presented from two loudspeakers positioned directly on the sides of the two ears (note, speech was always presented from the speaker in front of the subject). Again the best results were obtained with 2-band compression hearing aid. The average SRT obtained with the 2-band compression aid was 2.5 dB lower than with the high-fidelity linear hearing aid, which in turn was 1.3 dB lower than with the subject's own, or no hearing aid. Listening with 2 aids produced a lower SRT than listening with one aid. SRT was further reduced when the signal and noise were spatially separated. This improvement of SRT was greatest for the 2-band compression system. This study also found significant inter-subject variability. While some subjects showed only a small improvement, others showed substantial improvement. Laurence et al. also obtained subjective ratings of the different hearing aids. The 2-band compression system obtained the best ratings. Subjects rated the 2-band compression hearing aid as providing the largest improvement of speech intelligibility in difficult listening conditions.

The most salient result of the Laurence et al. study was improved intelligibility in noise with the 2-band compression hearing aid. This result is somewhat surprising, since most other work on compression hearing aids has shown that compression does not improve speech intelligibility in noise in comparison to linear amplification, and often makes it worse (e.g., Lippmann et al., 1981).

Moore, Laurence, and White (1985) re-examined the benefit of binaural 2-band compression hearing aids in noise. Using the 2-band compression hearing aid reported by Laurence et al. (1983), SRT in speech-shaped noise was obtained from 8 subjects with bilateral moderate-to-severe sensorineural hearing loss. The fitting procedure for the 2-band compression system was the same as that used by Laurence et al. (1983). Noise was presented at 65 and 75 dB SPL. Both speech and noise were presented from the same loudspeaker positioned directly in front of the subject (coincidental condition). Scores for the 2-band compression system were compared with that for unaided listening. On average, there was 4.4 dB improvement of SRT in noise with 2-band compression hearing aid over unaided listening. They also measured the dynamic range of hearing for speech, that is the range from SRT in quiet to the highest comfortable listening level for speech. This range was found to be substantially larger for the 2-band compression hearing aid. Thus, this study too found 2-band compression system to be beneficial in background noise. However, there is an obvious drawback in the choice of the control condition in this study. The subjects in this study all had greater hearing loss at high frequencies than at low frequencies. Therefore, part of the improvement can probably be attributed to the aids making audible high-frequency components of speech that would have been inaudible in the unaided condition.

Moore and Glasberg (1986) compared their 2-band syllabic compression system with a single-band AGC system and unaided listening. The single-band AGC system also had 2 bands, but these were used only for frequency response shaping. Eight subjects with bilateral moderate sensorineural hearing losses were used. Speech and noise were presented via a single loud-speaker directly in front of the subject. Noise levels of 60 and 75 dB SPL were used. The average SRT for the 2-band compression system was significantly lower than that for both single-band AGC and the unaided condition. In 75 dB SPL noise, the SRT for the 2-band compression system was, on average, 2.9 dB lower than that for the single-band AGC system. This finding suggests that a 2-band syllabic compression system (which also incorporated input slow-acting AGC) is better than a simple AGC system. That is, compression of both the short-term and long-term intensity variations of speech yields superior performance to that provided by compression of long-term intensity variations alone.

The series of studies by Moore and his colleagues have shown that 2-band compression hearing aids are superior to single-band compression and linear hearing aids for people with moderately-severe hearing losses. The dynamic range of hearing for individuals with profoundly deaf is much smaller than that of individuals with moderately-severe hearing loss, and therefore the need for reduction of short-term dynamic intensity range of speech should be the greater (Boothroyd, Springer, Smith, and Schulman, 1988).

Boothroyd et al. (1988) measured the effect of 2-band compression limiting on speech perception by adolescents with profoundly impaired sensorineural hearing loss. In this system, the input was first fed to a slow-acting AGC to control for variations in the overall level of speech. The output of the AGC was then high-frequency emphasized to compensate for the high-frequency roll-off of the speech spectrum. Following this, the signal was split into 2 bands: a low-frequency band (70-1000 Hz), and a high-frequency band (1000-7000 Hz). Within each band, there was a gain control and fast-acting AGC. The output from the two bands was then mixed (here, frequency response between the two bands could also be adjusted). Following the mixer, a power amplifier was used to control the final output level. The system could function as a compression aid or a linear aid.

A 3-interval forced choice test of speech contrast perception (THRIFT) was administered to a group of 9 subjects. Testing was conducted monaurally. The 2-band compression system was compared to linear amplification alone. Note that the 2-band compression system had a built in high-frequency emphasis. Results revealed that for 8 of the 9 subjects, compression reduced performance. For 1 subject, however, performance was better with the 2-band compression system. This study suggests that though profoundly impaired subjects may seem to require syllabic compression, amplitude compression may not be beneficial to these subjects. Boothroyd et al. opined that profoundly impaired subjects generally have very poor frequency resolution. They tend to rely more on temporal cues for speech perception. Compression hearing aids, by distorting the temporal structure of the speech signal, may negate any benefit that may be otherwise expected.

Most commonly, subjects with sensorineural hearing loss tend to exhibit narrower dynamic range at high frequencies than at low frequencies. Sometimes the dynamic range at high frequencies can be as low as a few dB (Boothroyd et al., 1988). In contrast, the dynamic range of hearing at low frequencies is relatively large, and it is not uncommon to see a dynamic range of 30 dB or greater in some subjects with sensorineural hearing loss. For these subjects, then, it could be argued that syllabic compression is warranted only in the high-frequency region.

Moore and Glasberg (1988) compared their original 2-band compression hearing aid with a new 2-band system. In this newer system, a new dual front-end AGC system was incorporated to compensate for variation in the overall level of speech and also to deal effectively with sudden intense transients. Following the operation of the dual front-end AGC, the signal was split into two frequency bands. Syllabic compression could be applied in either or both of the 2 bands.

The following conditions were evaluated.

1) Dual front-end AGC alone
2) Dual front-end AGC with compression in high frequencies alone (Mark II)
3) Mark II + low-frequency compression
4) Original 2 band compression (Mark I)
5) Linear amplification

Subjects with moderately-severe hearing losses participated in this study. SRT was evaluated in quiet and in noise. The noise level was 70 dB SPL. In quiet, the SRTs for the different aided conditions were generally similar, except that SRTs for the linear condition tended to be slightly higher than those for the compression conditions. The measurement of SRTs in noise showed that the slow-acting AGC alone with syllabic compression in the high-frequency bands gave a significant improvement over linear and front-end AGC alone. The addition of a low-frequency band did not usually result in an improvement. Thus, this study shows that for some subjects, low-frequency compression is unnecessary. However, the dual-front AGC was not found to be superior to the original front-end AGC. It should be mentioned that in this study the noise had a steady overall level. The type of dual-end AGC used in this study would be expected to be of benefit when there are fast intensity variations in the background noise.

The dual front-end AGC may be beneficial in situations where the noise is fluctuating such as general background noise plus door slamming, traffic noise etc. Moore, Glasberg, and Stone (1991) conducted an experiment to evaluate and optimize the dual front-end AGC. In this experiment, the release time of the fast-acting AGC and the attack time of the slow-acting AGC were varied. The following release/attack times were used: 40/70, 80/50, 80/325, 150/325 ms. The attack time of the fast-acting AGC and release time of the slow-acting AGC were kept constant at 5 & 7500 ms, respectively. The speech intelligibility scores obtained with this system were compared to a linear hearing aid and an adaptive compression hearing aid. The constants of the adaptive compression hearing aid varied automatically as a function of the duration of the signal. The speech stimuli consisted of sentences that were presented at 75 dB RMS level. At the beginning of each sentence an intense transient was included. The sentences were presented in quiet, background of speech-shaped noise, and single-talker speech (time reversed). For the 2 noise backgrounds, a S/N ratio of 5 dB was maintained. The dependent variable measured was percent key words recognized.

The optimal release time of the fast-acting AGC was 80 to 150 ms. The optimal attack time of the slow-acting AGC was 150 to 325 ms. When the speech signal containing transients was presented either in quiet or in speech shaped noise background, performance with the dual front-end AGC was better than that of the linear hearing aid and adaptive compression hearing aid. When time-reversed speech was used as background noise, dual front-end AGC and adaptive compression hearing aids performed comparably. Further, when the speech levels were raised over a 30 dB range, performance with both compression hearing aids remained essentially constant, but performance with the linear hearing aid worsened markedly at the lower levels. Thus, as expected, the 2-band compression system with dual front-end AGC was superior to a linear aid as well as an adaptive compression aid when the background noise had significant temporal variations.

Moore, Johnson, Clark, and Pulvinage (1992) evaluated the Resound in-the-ear hearing aid that applied fast-acting full dynamic range compression independently in two frequency bands. In this system, the input was fed to a compression limiter with a high threshold (85 dB SPL) and compression ratio (20:1). The attack and release times of the limiter were 1 ms/ 40 ms. The limiter provides protection against brief intense sounds. Following the limiter, the signal was split in two bands. The crossover frequency between the two bands could be adjusted depending on a weighted average threshold in the 500 to 4000 Hz range. In each band, the compression threshold was set at 45 dB SPL. The gain and compression ratios were calculated by weighting the average shapes of the loudness intensity function at 500, 1000, 2000, and 4000 Hz. The output of the 2 bands were mixed and delivered through the receiver. This hearing aid could also be programmed to function as a linear hearing aid with compression limiter for levels 85 dB SPL and above. In the Moore et al. study, the frequency response of the aid in the linear mode was adjusted to provide the frequency response that the compression aid would provide for a 65 dB SPL input.

The 2-band compression system was compared to the 2-band linear system, the subjects' own hearing aids and unaided listening. A total of 20 subjects with mild-to-moderately severe hearing losses participated in this study. Percent keywords recognized in sentences were measured in quiet as well as in 12-talker babble background noise. Performance was evaluated for 50, 65, and 80 dB SPL speech input levels.

In quiet, as the speech levels were decreased from 80 dB to 55 dB, performance remained essentially unchanged for 2-band compression. For the other 3 conditions, performance deteriorated as the input was lowered. For 2-band linear amplification and the subjects' own hearing aid, performance was comparable to the 2-band compression aid for the 80 and 65 dB SPL input levels. At the 50 dB SPL input levels, the performance with 2-band linear amplification and the subjects' own aid was markedly lower. Unaided performance was comparable to the other 3 conditions only for the highest input level of 85 dB SPL.

Moore et al. also evaluated the effect of independent compression at the two ears over binaural linear amplification. They presented the speech signal from a loud speaker directly in front of the subject. Noise was presented in 2 different ways; in the first way, noise was presented only with the speech from the loud speaker directly in front of the speaker (coincidental stimuli). In the second way, the noise was presented from loud speakers situated on either side of the two ears (spatially separated). As expected, performance was better when speech and noise were spatially separated, for both compression and linear amplification conditions. More importantly, binaural compression provided SRTs comparable to that of binaural linear hearing aids, for coincidental and spatially separated conditions. Moore et al. concluded that independent compression in the binaural aids does not necessarily affect speech recognition ability.

Moore, Lynch & Stone (1992) investigated the effect on SRT, in quiet and in noise, of varying the low-level gain (and hence the compression ratio) of a simulated Resound 2-band hearing aid. Not surprisingly, SRT in quiet decreased as the low-level gain and compression ratios were increased. SRT in noise was largely unaffected, but at the highest noise level (75 dB SPL), an increase in low-level gain in the low-frequency band caused a 2 dB worsening of SRT. Thus, Moore et al. concluded that excessive gain for low-level signals is unwarranted.

Benson, Clark, and Johnson (1992) compared the Resound hearing aid with the subject's own aids for 18 subjects with mild to severe sensorineural hearing loss. At each aid's used volume control setting, the Resound aid resulted in significantly lower aided thresholds, higher LDLs and consequently wider dynamic ranges. It also resulted in significantly higher speech identification scores in quiet and higher rating in a 10-item questionnaire. Overall, the results were strongly in favor of the Resound aid.

Thus, these studies have shown that 2-band compression is superior to unaided listening, high-frequency emphasis linear amplification, and single-band compression. Among the benefits reported for the 2-band compression are superior performance in both quiet and noise, especially when the input level of the speech signal is reduced.

A number of issues have been raised to account for the conflicting findings with syllabic compression. These issues will now be discussed.

2.4.4. Critique on negative effects of multi-band syllabic compression:

Plomp (1988) predicted that the effect of multi-band compression would be detrimental as the number of compression bands is increased and as the time constants of the compressors are decreased. The basis of his argument relied on the Speech Transmission Index (STI) and the Modulation Transfer Function (MTF) concepts. According to these concepts, preservation of envelope fluctuations of the target speech signal is a predictor of speech intelligibility. Plomp noted that noise and reverberation reduce MTF and since presence of noise and reverberation are known to affect speech intelligibility, there is a strong relationship between MTF and speech intelligibility. He opined that any reduction of MTF would result in poor speech intelligibility. He went on to show that both short time constants as well as increasing the number of processing bands reduce MTF and therefore would cause poorer speech intelligibility.

On the other hand, Villchur (1989) argued that the concept of STI does not hold good for amplitude compression. Villchur pointed out that noise and reverberation reduce the MTF by filling in the troughs of signal-amplitude modulations. Speech intelligibility is reduced along with the MTF, not because the MTF is reduced, but because the weaker elements of the speech signal are masked and the information they carry is lost. In contrast, amplitude compression reduces the MTF by bringing the troughs and peaks of the speech modulations closer together. The weak speech elements that are lost when MTF is reduced by noise are enhanced when the MTF is reduced by compression. In reply to this remark, Plomp (1989) suggested that this issue can be resolved by studying speech intelligibility of amplitude compressed speech as a function of increasing number of compression bands and compression ratios. Plomp (1994) cited an unpublished study of Van Dijkhuizen. In this study, percent word recognition for sentences in noise was measured as a function of number of bands (1,2,4,8,16) and compression ratios (2:1, 4:1, ∞:1) in a group of 16 moderately cochlear impaired listeners. As the number of bands was increased, % words recognized decreased and this decrease was greater for higher compression ratios. For compression ratio of 2:1, 2- and 4-band compression did not degrade performance substantially over linear amplification. Thus, Plomp concluded that too many compression bands and too large compression ratios have deleterious effects on speech intelligibility.

Moore (1987), and Levitt and Neuman (1991) offered different explanations to account for poor results with amplitude compression. These investigators opined that compression amplification smoothes the short-term speech spectrum and that this smoothing of the spectrum increases with increasing the number of compression bands. Since picking of salient peaks of the speech spectrum is important for speech intelligibility, smoothing of the speech spectrum by amplitude compression may account for poorer speech intelligibility. Moore (1987) suggested that by limiting the number of compression bands smoothing of the speech spectrum could be contained. Levitt and Neuman (1991) suggested orthogonal polynomial compression in which a family of orthogonal polynomials first approximates the speech spectrum. The coefficients of these polynomials could be compressed independently.

Villchur (1983) suggested that methodological issues might have caused poor results with compression amplification. He contended that the validity of the test of amplitude compression effectiveness may be compromised by 1) the use of speech test material of limited dynamic range, 2) the use of high-frequency emphasis for the uncompressed reference signals in an amount suitable for limited dynamic range test material, but creates unacceptable stridency in real-life speech, and 3) the use of subjects whose residual dynamic range of hearing is not restricted enough to be a significant handicap for speech reception.

While all of the above issues may have a bearing on the negative findings with multi-band compression, a more fundamental issue of the effectiveness of syllabic compression in reducing the dynamic intensity range of speech should also be considered. It should be pointed out that most compression hearing aids are traditionally validated using steady-state signals such as pure tones. While the effect of compression on such steady-state signals can be easily predicted, its response to a dynamic signal such as speech is less predictable. Some data suggests that syllabic compression does not always produce the desired effect.

Krieg (1980) evaluated the effectiveness of the 16-band compression system used by Lippmann et al. (1981). Effective compression ratio was measured using input/output functions for pure tones and by comparison of the cumulative amplitude distributions of the uncompressed and compressed CVC word stimuli. As expected, Krieg's results revealed that the desired compression ratio was obtained for input/output functions using pure tones. Comparison of the cumulative amplitude distributions of the speech stimuli revealed, however, that the obtained compression ratios fell short of the desired compression ratios. Thus, this study points out that the traditional input/output functions using steady-state signals do not reveal the effectiveness of compression of dynamic signals such as speech. The author suggested that the time constants of the compression, specifically the release time, are the most likely contributor to the disparity between the desired and obtained compression ratios.

Kuk (1996) illustrated that as the release time is increased, the effective compression ratio decreases. Stone and Moore (1992) stated that the effective compression ratio is influenced by the relation of the release time of the compressor to the magnitude of the inter-syllabic duration. Generally, If the release time of the compressor equals the duration of the syllable, then the obtained compression ratio would approximate that predicted. If the release time increases relative to the duration of the syllable, the effective compression ratio would be less than that desired. From the above logic, therefore, it can be speculated that a look-forward compression scheme that is not constrained by the attack/release times would yield a closer match between the desired and obtained compression ratios. The effectiveness of such a scheme was evaluated in the present study.

Another potential factor that could undermine the effective compression ratio is the temporal integration time of the cumulative amplitude distribution. It was noted previously in this chapter that the short-term dynamic intensity range of speech is ideally measured using a temporal integration time corresponding to the temporal integration time of the ear, which is roughly about 100 ms. Therefore, if a cumulative amplitude distribution measure is used as an indicator of the effectiveness of compression, then the integration time of the measure should correspond to the integration time of the ear. Such a measure would give a closer estimate of the perceived dynamic range of speech. Souza and Turner (1996) examined the effectiveness of a single-band syllabic compressor (attack/release time of 8/15 ms, and compression ratio of 5:1) as revealed by 3 different measures. They obtained input/output functions for pure tones, and compared temporal waveforms of uncompressed and compressed speech waveforms and cumulative amplitude histograms of unprocessed and compressed speech stimuli that were obtained using a 128 ms temporal integration time. While the desired compression ratio was achieved for input/output functions and comparison of temporal waveforms, the cumulative amplitude distribution revealed that the compression was fully effective for only a minor portion of the distribution. Souza and Turner postulated that since syllabic compression controls only relatively fine intensity variations in the speech signal (in this case not more than 15 ms), such variations when averaged over a long duration are bound to show decreased effectiveness.

The theoretical implications of the above finding would suggest that, on one hand, the detection time of the compressor should be long enough to be comparable to the temporal integration time of the human ear, so that the desired compression ratio is obtained. Unfortunately, this would require us to lengthen the time constants of the compressor. When the time constants of the compressor are larger than the average duration of a speech syllable, which is approximately 40 ms, the compressor would no longer be a syllabic compressor. On the other hand, it can be argued that the applied compression may be increased to produce a lesser, but desired effective compression ratio. Unfortunately, again, past research on syllabic compression has shown that the human ear is limited in terms of the extent of compression it can tolerate. For example, Verschuure et al. (1994) has shown that the speech recognition scores decrease as the compression ratio is increased. This paradoxical situation should be viewed as an imposed natural limitation. Perhaps we should be more realistic in our expectation of the benefits of compression.

The acoustical measurement of the effectiveness of compression, even when obtained using speech stimuli has limited face validity. Specifically, the acoustical measurements of speech do not necessarily indicate the perceptual outcome. As was pointed out earlier in this chapter, while both acoustical and perceptual measures of the dynamic intensity range of speech have been reported, there is no single comprehensive study, which relate these two measures. In the present study, PI function was used as a perceptual validation procedure for the measurement of the effective compression ratio. Using this procedure, the width or the slopes of the PI function for compressed and uncompressed speech stimuli could be compared to provide an estimate of the effective compression ratio.

To summarize, compression has been used in hearing aids to provide protection against discomfort for loud sounds, to compensate for the reduced dynamic range of hearing, to compensate for loudness recruitment, and to reduce the deleterious effects of background noise. A number of studies have been conducted to assess the benefits of syllabic compression as a means of reducing the short-term dynamic range of speech. In general, these studies have provided conflicting results.

Prior to the 1960's, studies on syllabic compression focused mainly on single-band compression. The results of these studies were generally negative, probably because of factors related to inappropriate selection of compression parameters, system-induced distortions, poor design, and inadequacy of single-band compression in reducing the dynamic range of speech. During the 1970's, studies with multi-band compression yielded conflicting results, probably because of inappropriate selection of the reference condition. Indeed, studies during the late 1970's demonstrated that effectiveness of multi-band compression is negligible when compared with optimal high-frequency emphasis linear amplification. During the early to mid 1980's, studies on multi-band compression used as many as 16 bands. Results of these studies again showed no benefit of compression over carefully chosen linear amplification condition. However, these studies did point out that compression is superior when there are variations in the overall speech input levels. The studies that showed positive results with compression amplification used only a few bands. Such studies with moderate to moderately-severe hearing loss individuals showed syllabic compression to be of benefit over linear amplification. Among the benefits reported was superior intelligibility as the overall level is varied, in quiet, and in noise. In contrast, studies with profound hearing loss were limited. One study with this population revealed that compression was detrimental to speech recognition when compared with linear amplification.

A number of issues have been raised to account for the conflicting results which include: the use of subjects who do not need compression, a failure to match compression parameters to the characteristics of the speech and/or the experimental subjects, deleterious spectral and temporal changes accompanying syllabic compression, acoustical outcome evaluation with steady-state stimuli rather than dynamic speech-like stimuli, and perceptual outcome evaluation involving testing at high sensation levels at which compression offers little or no advantage.

2.5. PURPOSE OF THE PRESENT STUDY:

The purpose of the present study was to evaluate the potential benefits of wide dynamic range compression amplification for persons with sensorineural hearing loss. Specifically, the effectiveness of single-band and 2-band fast-release wide dynamic range compression in reducing the short-term dynamic intensity range of speech was evaluated. Their effectiveness was compared against no-processing and simple high-frequency emphasis alone conditions. Effectiveness was measured using both acoustical and perceptual measures. Both measures incorporated speech as the stimuli. The acoustical evaluation involved the measurement of RMS cumulative amplitude distributions and dynamic input/output functions. For, the high-frequency emphasis alone condition, the extent of reduction of the RMS cumulative amplitude distribution was evaluated. For the compression conditions, the extent to which the obtained input/output functions matched to the target was evaluated. The perceptual evaluation involved the measurement of PI functions using normal-hearing individuals listening in pink-noise. The effect of amplitude compression on the width and the slope of the PI function was evaluated to provide an estimate of upper limits of possible benefits that can be achieved when such processing schemes are tried on individuals with sensorineural hearing loss.

2.6. RESEARCH QUESTIONS:

1. To what extent does high-frequency emphasis alone reduce the acoustic and perceptual dynamic range of speech?
2. When fast-release single-band compression is added to the pre-compression high-frequency emphasis, to what extent do the changes in the acoustic and perceptual dynamic range of speech match those predicted from static input/output functions?
3. When fast-release 2-band compression is added to the pre-compression high-frequency emphasis, to what extent do the changes in the acoustic and perceptual dynamic range of speech match those predicted from static input/output functions?

CHAPTER III: SIGNAL PROCESSING

3.1: PURPOSE:

To create digital audio CVC word stimuli for the following 4 conditions: a) no-processing, b) high-frequency-emphasis-alone, c) pre-compression high-frequency-emphasis plus fast-release single-band compression (compression ratio of 3:1), and d) pre-compression high-frequency-emphasis plus fast-release 2-band compression (compression ratio of 3:1 in both bands).

3.2: SPEECH MATERIAL:

Test material included digitized recordings (female talker) of AB iso-phonemic word lists. The AB iso-phonemic word lists consisted of 20 lists of 10 Consonant-Vowel-Consonant (CVC) words each (see Appendix A for the AB word lists)[1]. The same 10 vowels and 20 consonants appeared in each list. These are the vowels and consonants that occur most frequently in English CVC words. There were no constraints on word frequency or word familiarity, but each of the 200 words appeared only once in the complete set (Boothroyd, 1969).

Digital audio recordings of the AB word stimuli spoken by a female talker were made in an IAC, sound attenuating enclosure. The microphone had a uniform frequency response (+/- 1 dB) from 20 to 20,000 Hz. It was attached to a wind shield and was placed 4 inches from the talker's mouth at 0 degrees azimuth. The recorded words were digitized with a sampling rate of 22050 Hz

---

[1] The original AB word lists consisted of 15 lists. An additional 5 lists were created by the author for this and other experiments. Lists 16-20 in Appendix A are the newly added lists.

and 16-bit quantization. Each stimulus was provided with a 200 ms pre-onset and post-offset silent interval.

In addition to the AB word lists, NU 6 words were used in the perceptual validation experiments to allow subjects to become familiarized with the test procedure. The NU 6 words were presented live by the examiner. Only NU 6 words not appearing in the AB word lists were used for familiarization.

3.3: TEST CONDITIONS:

The following 4 conditions were evaluated in this study:

1. Naturally Spoken Unprocessed (referred to henceforth as No-Processing),
2. High-Frequency(HF)-emphasis-alone,
3. Pre-compression HF-emphasis plus fast-release Single-Band Compression, and
4. Pre-compression HF-emphasis plus fast-release 2-Band Compression

The rationale for the selection of the above 4 conditions is as follows. The no-processing condition was selected to measure the talker-specific short-term dynamic intensity range of the speech. Generally, the average peak intensity level of the HF band of the speech signal is less than that of the low-frequency (LF) band. Thus, to some extent a simple HF-emphasis alone should be able to reduce the short-term dynamic range of speech. Thus, an HF-emphasis-alone condition was also selected. Any further reduction of the short-term dynamic range of speech would require fast-release compression. Two forms of fast-release compression, single-band compression and 2-band compression, were included in this study to evaluate if 2-band compression is more effective in reducing the dynamic range of speech than single-band compression. While single-band systems can reduce the intensity variations among successive syllables, multiband systems can, in addition, reduce the intensity variations across frequency at a given time. Note that both single-band and 2-band compression were applied to the HF emphasized signal. It was hypothesized here that as we move from no-processing, to HF-emphasis-alone, to pre-compression HF-emphasis plus fast-release single-band compression, to pre-compression HF-emphasis plus fast-release 2-band compression, compression of the short-term dynamic intensity range of speech would progressively increase both acoustically and perceptually.

3.4: PREPARATION OF TEST STIMULI:

3.4.1: Software:

The digitized AB words were processed to obtain the stimuli for the 4 conditions for this study. An array processing software package, DaDisp from DSP corporation, was used for this purpose. (See Appendix B for the details of the procedures used for signal processing).

3.4.2: Signal processing for no-processing condition:

Each stimulus was digitally filtered through a high-pass filter to remove the LF background noise below 300 Hz found in the original digital recordings. The high-pass filter had a cut-off frequency of 300 Hz, stop-band edge of 1 Hz, 3 dB ripple in the pass-band, and a 50 dB pass-band attenuation.

3.4.3: Signal processing for HF-emphasis-alone condition:

The goal of signal processing for HF-emphasis-alone condition was to equate the 90th percentiles of the cumulative RMS amplitude distributions of the low- and high-frequency bands of the unprocessed speech stimuli. The following steps were taken to accomplish this goal.

1) Filtering: Each stimulus of the no-processing condition was filtered through a low-pass and a high-pass filter. The low-pass filter had a cut-off frequency of 950 Hz, stop-band edge of 1850 Hz, 1 dB ripple in pass-band, and a 100 dB pass-band attenuation. The high-pass filter had a cut-off frequency of 1250 Hz, stop-band edge of 500 Hz, 1 dB ripple in pass-band, and a 100 dB pass-band attenuation. The crossover frequency between the bands was 1100 Hz. Figure 3.1 shows the transfer functions of the 2 filters. The characteristics of the two filters were chosen such that, generally, for a female talker, the first formant would fall in the LF band and the second formant would fall in the HF band. Also, when the outputs of the two filters were recombined, these filter characteristics ensured minimal distortion of the spectrum in the region of the crossover frequency, as shown in Figure 3.1.

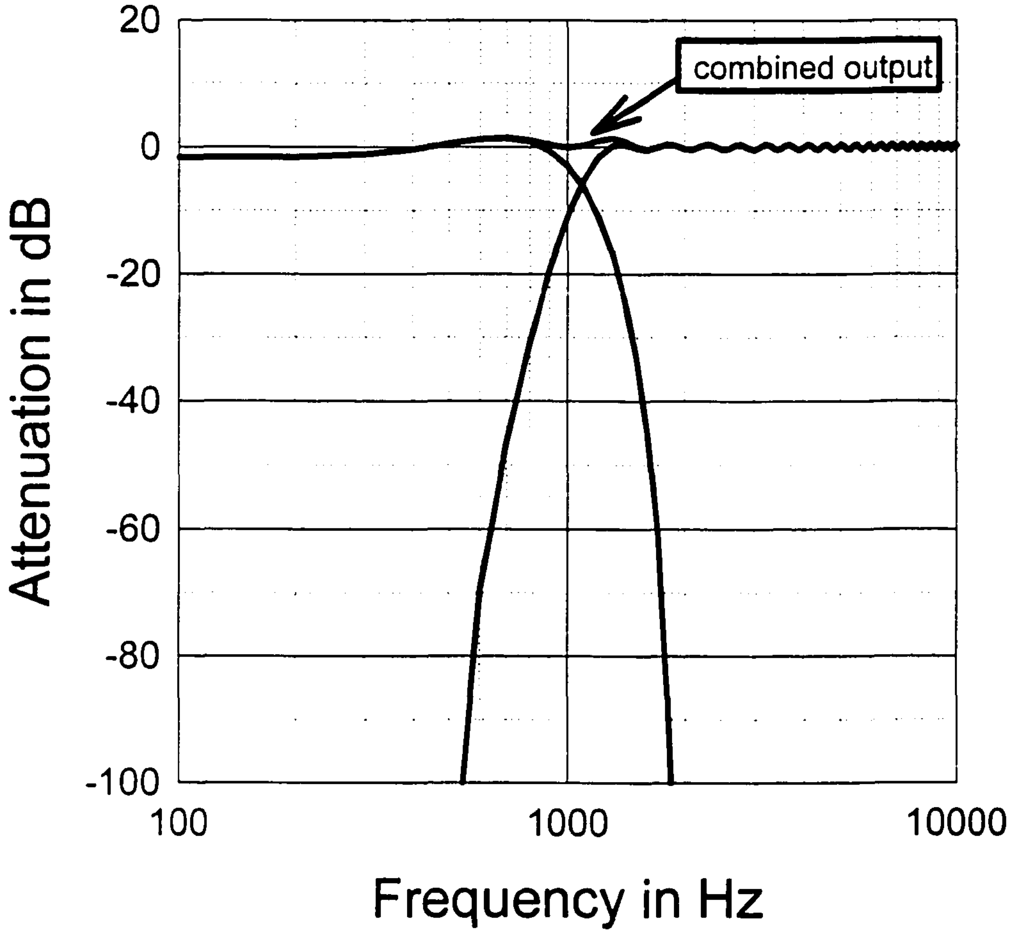

Figure 3.1: Attenuation as a function of frequency for the digital low-pass and high-pass filters used in this study. Also shown is the effect of combining the filter outputs.

2) Amplification of the HF band: Cumulative RMS amplitude distributions of the LF and HF bands of the no-processing condition were obtained, using a temporal integration time of 125 ms, to determine the amount of HF emphasis needed to equate the 90th percentile distributions of the 2 bands (see Appendix C for the details of the measurement procedure used for obtaining the cumulative amplitude distributions). Figure 3.2 shows the mean RMS amplitude distributions of the LF and HF bands of the entire 200 word stimuli of the no-processing condition.

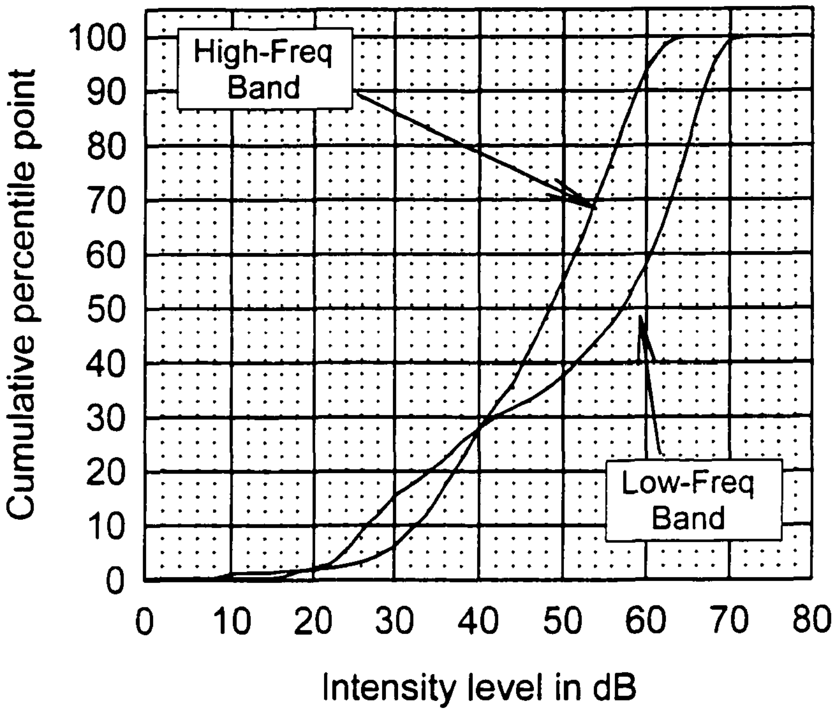

Figure 3.2: Mean cumulative amplitude distributions of the LF and HF bands of the no-processing condition obtained using a temporal integration time of 125 ms.

As the figure reveals, at the 90th percentile level, the HF band is about 8 dB lower than the LF band. Thus, the output of the high-pass filter was amplified by a constant 8 dB. Note that this processing did not intend to equate the peak levels of the 2 bands on an individual word basis. Rather, the 8 dB HF boost was intended to equate the 90th percentile of the mean cumulative amplitude distributions of the LF and HF bands of the entire 200 word stimuli of the no-processing condition.

3) Combining the LF and the emphasized HF bands: The output of the LF band was left unaltered. The unaltered low-pass filtered signal was combined with the amplified high-pass filtered signal to obtain the HF-emphasis-alone signal.

3.4.4: Signal processing for pre-compression HF-emphasis plus fast-release single-band compression condition:

An off-line digital look-forward fast-release single-band compression with a compression ratio of 3:1 was applied to the HF emphasized stimulus. The goals of this condition were: 1) to compress as much of the short-term dynamic range of speech as possible without allowing the low level ambient noise to be amplified, 2) to compress the intensity levels of individual word stimuli as well as inter-word level differences. The following steps were taken to realize this goal.

1) Choice of compression threshold: To aid in the choice of the compression threshold, the compression threshold of the processing algorithm was varied and listening checks were performed to locate a lowest compression threshold that did not result in a perceivable increase of ambient noise in the processed signal. Based on the listening checks, a compression threshold of 46 dB (in digital units) was chosen. This threshold was approximately 30 dB below the peak RMS level of cumulative amplitude distribution of the recorded material. To determine the dB extent of possible reduction of the short-term dynamic range of speech stimuli, a cumulative RMS amplitude distribution of the full band of the entire 200 word stimuli of the HF-emphasis-alone condition was obtained in the manner described in Appendix C. Figure 3.3 shows the cumulative RMS amplitude distribution of the full band for the HF-emphasis-alone condition.

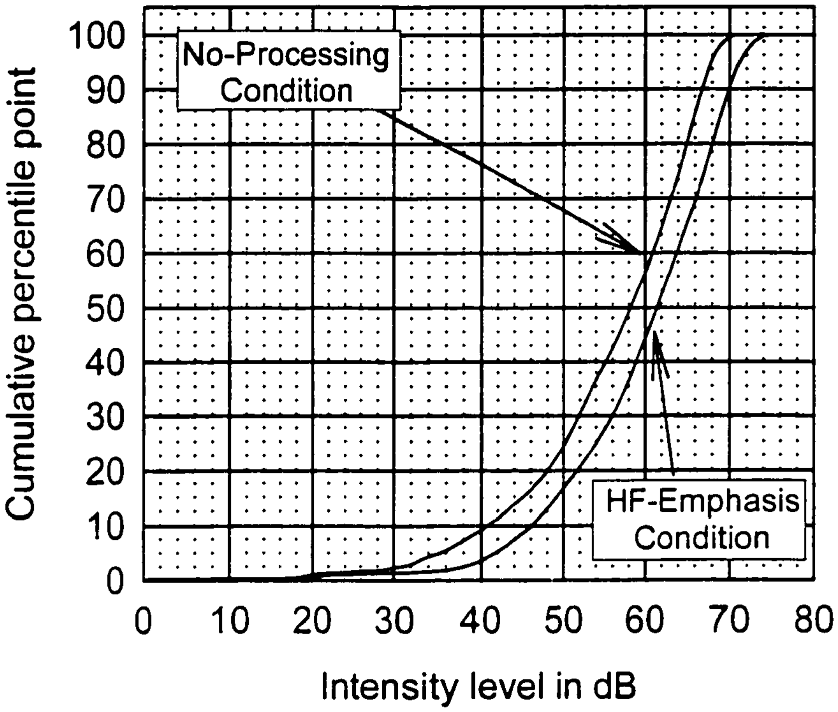

Figure 3.3: Mean cumulative amplitude distributions of the full-band of no-processing and HF-emphasis-alone conditions, obtained using a 125 ms temporal integration time.

As Figure 3.3 reveals, the 100% peak point of the distribution was 75 dB (in digital units). Thus, the compression threshold of 46 dB aimed at compressing the top 29 dB of the mean cumulative amplitude distribution by a factor of 3:1. Note that the compression threshold was approximately at the 10th percentile of the distribution. Therefore, in terms of percentiles, this corresponds to a compression of the top 90 percent of the amplitude distribution.

2) Compression process: Figure 3.4 illustrates the compression processing scheme. The RMS amplitude envelope of the HF emphasized signal was extracted using a temporal window of 5 ms duration, each window overlapping the previous one by 2.5 ms. The HF emphasized signal was then divided by its amplitude envelope to obtain a unity RMS amplitude signal. A compression threshold of 46 dB (in digital units) was established. The amplitude envelope above the compression threshold was compressed by a factor of 3:1. The amplitude envelope below the threshold was left unaltered. The compressed envelope was then given a 20 dB boost (post-compression amplification). The purpose of the post-compression amplification was to restore the peak levels of the compressed signal. Finally, the unity RMS amplitude signal was multiplied with its compressed envelope to obtain the pre-compression HF-emphasis plus fast-release single-band compressed signal.

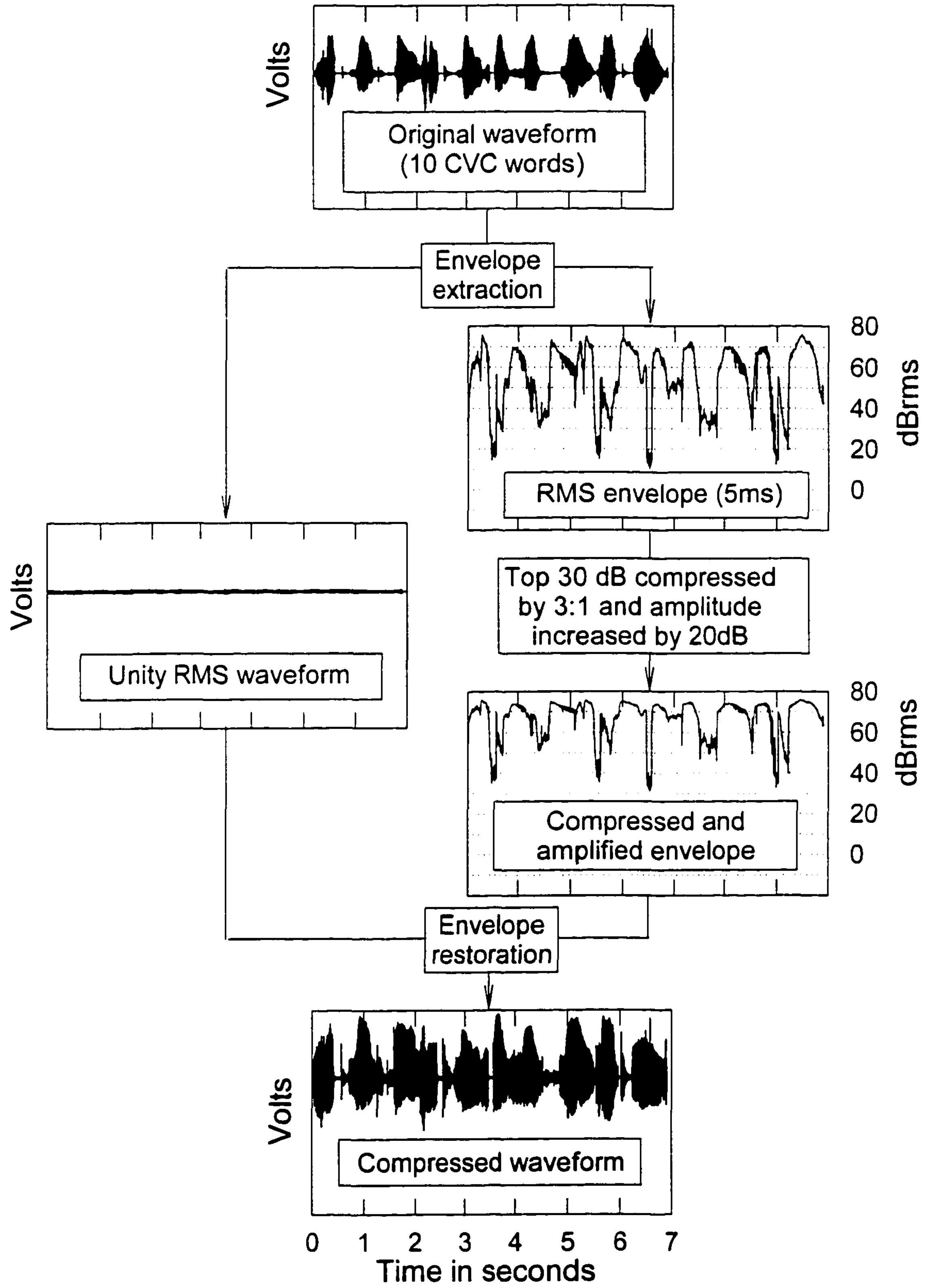

Figure 3.4: Illustration of the algorithm used to accomplish compression. Note the 20 dB post-compression boost, designed to equate the maximum RMS level for compressed and uncompressed signals.

3.4.5: Signal processing for pre-compression HF-emphasis plus fast-release 2-band compression condition:

An off-line digital look-forward fast-release 2-band compression with a compression ratio of 3:1 was applied to the HF emphasized stimulus. Within each band, the goals of this condition were: 1) to compress as much of the short-term dynamic range of speech as possible without allowing the low level ambient noise to be amplified, and 2) to compress the intensity levels of individual word stimuli as well as inter-word level differences. The following steps were taken to realize this goal.

1) **Filtering:** The original uncompressed stimulus was fed through a low-pass filter and a high-pass filter (filter characteristics were same as described in signal processing for HF-emphasis-alone condition).

2) **Compression Process:** Compression and post-compression amplification were applied to the output of the 2 filters in the same way as described in the signal processing for fast-release single-band compression section. An 8 dB emphasis was given to the high-pass filtered signal before compression was applied. A compression threshold of 46 dB and a compression ratio of 3:1 were used for both LF and HF bands. Note that the compression threshold of 46 dB was chosen to maintain uniformity between that of the single-band and 2-band compression conditions, and in the case of 2-band compression between the LF and HF bands. Figure 3.5 shows the cumulative amplitude distribution for the LF and HF bands for the HF-emphasis-alone condition.

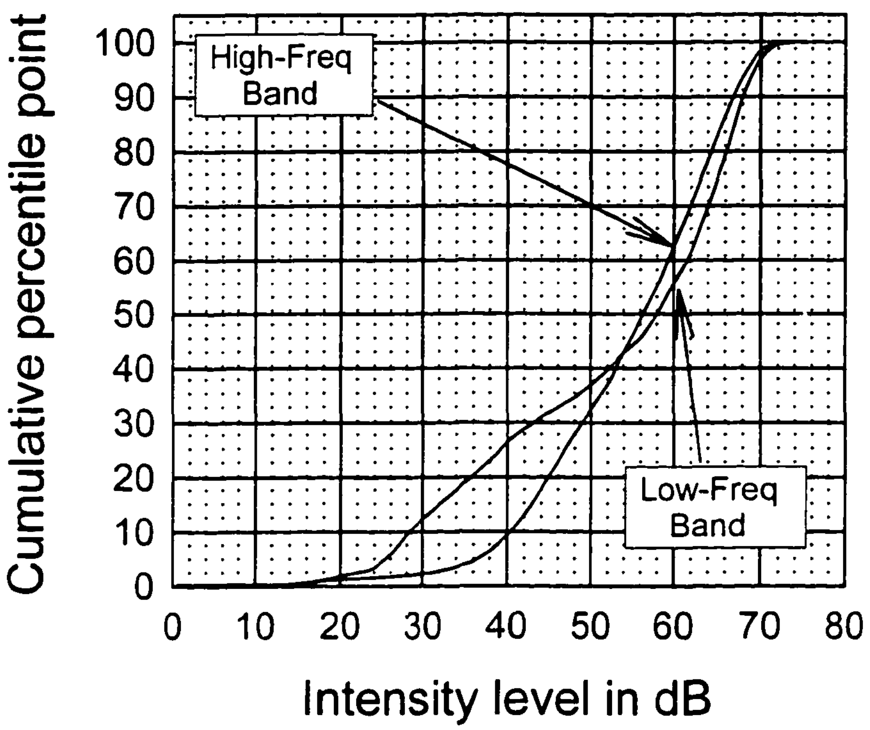

Figure 3.5: Mean cumulative amplitude distributions of the LF and HF bands for the HF-emphasis-alone condition obtained using a temporal integration time of 125 ms.

Figure 3.5 reveals that the 100% peak point of the distribution for the LF and HF bands for the HF-emphasis-alone condition was approximately 73 dB. Thus, the 2-band compression would result in compressing the top 27 dB of the cumulative amplitude distributions of the LF and HF bands. Figure 3.5 also reveals that the compression threshold was at the 33rd percentile for the LF band and at 23rd percentile for the HF Band. Therefore, the 2-band compression aimed at compressing the top 67 and 77 percent of the distributions for the LF and HF bands, respectively. The 2 compressed bands were then recombined to obtain the speech signal with pre-compression HF-emphasis plus fast-release 2-band compressed signal.

If we assume that the bulk of the useful information in speech is contained in the top 30 dB of the amplitude distribution, then the compression used here should have compressed the bulk of the useful information into a 10 dB range, theoretically reducing the range of the performance-intensity function from 30 dB to 10 dB.

CHAPTER IV: ACOUSTICAL EVALUATION

4.1: PUPROSE:

To evaluate how closely the changes in RMS amplitude distribution produced by the different processing schemes correspond with the target.

4.2: ISSUES:

A number of issues have been raised to account for the many negative findings with compression hearing aids. The present study focuses on the following issues pertaining to acoustical evaluation of compression:

1) The algorithm's output is predictable for steady-state pure tones. Its behavior with a dynamic signal such as speech is theoretically predictable, but much less easily.
2) The measured distribution of the RMS amplitude is affected by 3 factors -
   a) the frequency band over which energy is integrated,
   b) the duration of the temporal window over which energy is integrated, and
   c) the shape of the temporal window.

For the present study, data were collected for:

   a) 3 frequency bands -
      i) the LF band,
      ii) the HF band, and
      iii) the full band

Note that the filter characteristics were the same as that used for signal processing for HF-emphasis-alone, fast-release single-band compression, and fast-release 2-band compression schemes.

- b) two rectangular temporal windows -
  - i) 5 ms - corresponding with the integration time of the compression algorithm, and
  - ii) 125 ms - approximating the temporal integration time of the human ear.

4.3: SPECIFIC QUESTIONS:

1) For the HF-emphasis-alone condition,
   - a) does the 8 dB boost of the HF band equate the 90% peak levels of the cumulative amplitude distributions of the LF and HF bands?
   - b) does the HF-emphasis result in a reduction of the dynamic range of speech when measured using the full band?
2) For the fast-release compression conditions,
   - a) do the algorithms produce the target performance when assessed using the frequency band(s) and the temporal window used in the algorithms?
   - b) if target compression is attained for full band analysis using 1-band compression, what is the effect of measuring in the LF and HF bands separately?
   - c) if target compression is attained for the LF and HF bands of the 2-band compression, what is the effect of combining those bands?
   - d) if target compression is attained for 5 ms integration time, what is the effect of measuring over 125 ms?

4.4: METHODS:

4.4.1: Acoustical evaluation procedure for HF-emphasis-alone condition:

Mean cumulative RMS amplitude distributions of the LF band, HF band, and the full band of lists of consonant-vowel-consonant word stimuli of the "HF-emphasis-alone" condition were obtained using a 125 ms temporal integration window in the manner described Appendix C.

4.4.2: Acoustical evaluation procedure for fast-release compression conditions:

Dynamic input/output functions were obtained for both 5 ms and 125 ms temporal integration times using DaDisp in the following manner (see Appendix D for a detailed explanation of the procedure).

Within a DaDisp worksheet, all 10 words of AB word list 1 of a given condition were imported as signed integer/binary files. The pre-onset and post-offset silent intervals present in each stimulus were deleted. The extracted stimuli were then concatenated and filtered through the low-pass and the high-pass filters used in the signal processing. The amplitude distributions of the unfiltered (full band), low-pass filtered (LF band), and high-pass filtered (HF band) concatenated strings were then computed.

This process was carried out separately for the HF-emphasis, fast-release single-band compression, and fast-release 2-band compression conditions. To obtain the dynamic input/output functions for a specific frequency band as well as a specific temporal integration window for a given compression condition, an XY plot was drawn showing amplitude for the HF-emphasis-alone condition on the x-axis and the corresponding amplitude for the compression condition on the Y-axis. Each data point represents the intensity of a sample at a specific point in time.

To aid in quantifying the effectiveness of compression, the dynamic input/output functions were fitted with a predicted function of the form

$$y = -\log(10^{-(X + B)} + 10^{-((X - C)/D + (B + C))}) \quad \text{.................................} (1)$$

where,

y = output amplitude in dB

X = input amplitude in dB

B = post-compression gain = 20 dB

C = compression threshold = 46 dB

D = Compression ratio = 1/3

4.5: RESULTS:

4.5.1: Acoustical evaluation of HF-emphasis-alone condition:

1) Success of HF-emphasis in equating the 90th percentile points of the RMS cumulative distributions of LF and HF Bands:

Figures 3.2 and 3.5 of the previous chapter (pages 77 and 84, respectively) show the respective mean cumulative RMS amplitude distributions for the LF and HF bands of the no-processing and HF-emphasis-alone conditions obtained using a 125 ms temporal integration time. Inspection of Figure 3.2 reveals that for the no-processing condition, at the 90th percentile point of the distribution, the LF band is about 8 dB more intense than the HF band. Inspection of Figure 3.5 shows that the HF emphasis resulted in equating the LF and HF bands within about 1 dB at the 90th percentile point.

2) Effect of HF-emphasis on reducing the short-term dynamic range of speech as measured using the full band:

Figure 3.3 of the previous chapter (page 80) shows the full band cumulative RMS amplitude distributions for no-processing and HF-emphasis-alone conditions obtained using a temporal integration time of 125 ms. The dynamic range of speech from the 100th percentile point to the 10th percentile point is 30 dB for the no-processing condition and 29 dB for the HF-emphasis-alone condition. Therefore, in terms of dB extent of the distributions, there is only a minimal difference between the no-processing and HF-emphasis-alone conditions. Note, however, that at any given cumulative percentile point of the cumulative amplitude distribution, the HF-emphasis-alone condition is more intense than the no-processing condition. For most of the distribution, the HF-emphasis-alone is more intense by about 3 dB. These findings point to the limitation of the amplitude distribution alone as a means of assessing the distribution of importance in speech.

4.5.2: Acoustical evaluation of pre-compression HF-emphasis plus fast-release single-band compression condition:

1) Effectiveness of fast-release single-band compression for 5 ms temporal integration time:

Figure 4.1 shows the dynamic input/output functions for the full band, LF band and HF band of the fast-release single-band compression condition obtained using a temporal integration time of 5 ms. To quantify the accuracy of the processing schemes, the mean and standard deviation of the dB difference between the obtained and predicted amplitude values were calculated. Table 4.1 shows these values for this and other test conditions.

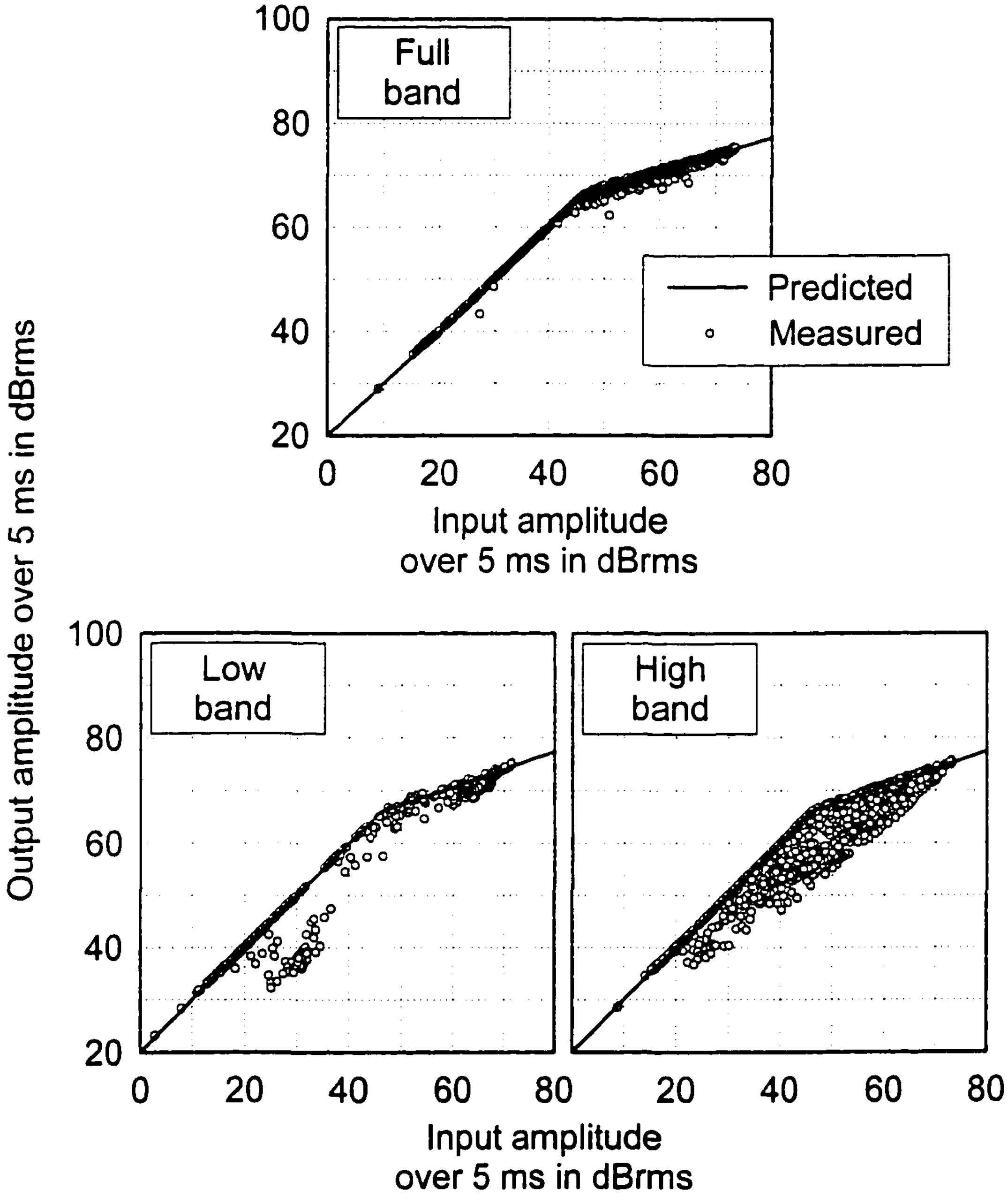

Figure 4.1: Measured and predicted input/output functions for single-band compression algorithm measured with 5 ms integration time.

Table 4.1: Mean and standard deviations of discrepancies in dB between the obtained and predicted input/output functions calculated over the entire amplitude distribution

	Single-band Compression				2-band Compression			
	5 ms		125 ms		5 ms		125 ms	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Full-Band	-0.06	0.48	-0.75	1.06	0.85	1.05	0.34	1.39
LF Band	-1.87	3.48	-2.3	2.68	0.13	0.68	-0.54	1.29
HF Band	-2.65	3.20	-3.33	2.83	-0.21	0.7	-1.06	1.33

Inspection of Figure 4.1 and Table 4.1 reveal that for the full band analysis, the obtained input/output function is in close agreement with that predicted. The mean and standard deviation of discrepancy between the obtained and predicted functions are in the order of only -0.06 dB and 0.48 dB, respectively. This finding suggests that when the acoustical confirmation procedure incorporates comparable measuring characteristics to that employed by the compression scheme, the target effect is easily achievable even for dynamic signals.

When analyzed separately in the LF and HF bands (Figure 4.1), the mean and standard deviation of discrepancy increase relative to that obtained for the full band data. For example, for the LF band, the mean and standard deviation of discrepancy are -1.87 dB and 3.48 dB, respectively. Similar large differences between the obtained and predicted input/output functions are also evident for the HF band. Therefore, it appears that the fast-release single-band compression scheme is less effective in reducing the intensity variations that often occur simultaneously in the low- and high-frequency bands of the speech signal.

2) Effectiveness of fast-release single-band compression for 125 ms temporal integration time:

Figure 4.2 shows the input/output functions for full band, LF band, and HF band for fast-release single-band compression condition obtained using a temporal integration time of 125 ms. Table 4.1 shows the mean and standard deviation of the discrepancy between obtained and predicted functions.

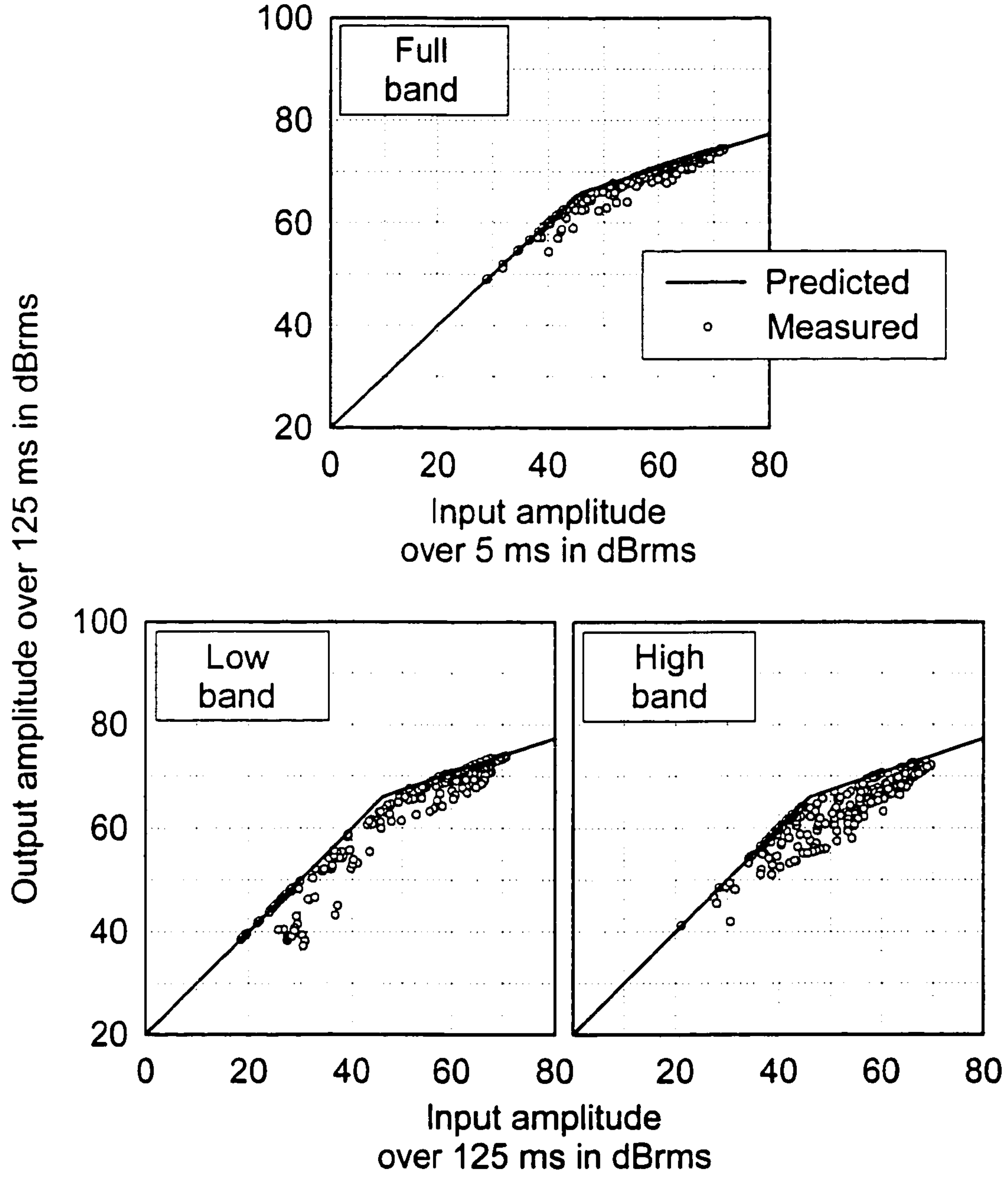

Figure 4.2: Measured and predicted input/output functions for single-band compression algorithm measured with 125 ms integration time.

Examination of Figure 4.2 and Table 4.1 reveals a closer agreement between the obtained and predicted dynamic input/output functions for the full band data, but not for the LF band and HF band data. This finding is similar to that found when analyzed using 5 ms temporal integration time.

Comparison of 5 ms and 125 ms integration time data reveals that the discrepancy between obtained and predicted input/output functions are much larger for the 125 ms temporal integration time than for 5 ms temporal integration time. For the full band analysis, the mean and standard deviation of the discrepancy increased to -0.75 dB and 1.06 dB. This finding indicates that as the integration time of the amplitude distribution measure increases, there is an overall decrease in effectiveness of the fast-release single-band compression scheme in meeting the target compression. The data also reveals that the mean discrepancy is even larger when analyzed separately in the LF and HF bands. This finding may be attributed to the fact that LF level is often governed by HF level - and vice versa - depending on the spectral location of the energy maximum at any given time.

4.5.3: Acoustical evaluation of pre-compression HF-emphasis plus fast-release 2-band compression condition:

1) Effectiveness of fast-release 2-band compression for 5 ms temporal integration time:

Figure 4.3 shows the dynamic input/output functions for the fast-release 2-band compression algorithm measured using a 5 ms temporal integration time. The mean and standard deviation of the discrepancy between obtained and predicted functions are shown in Table 4.1.

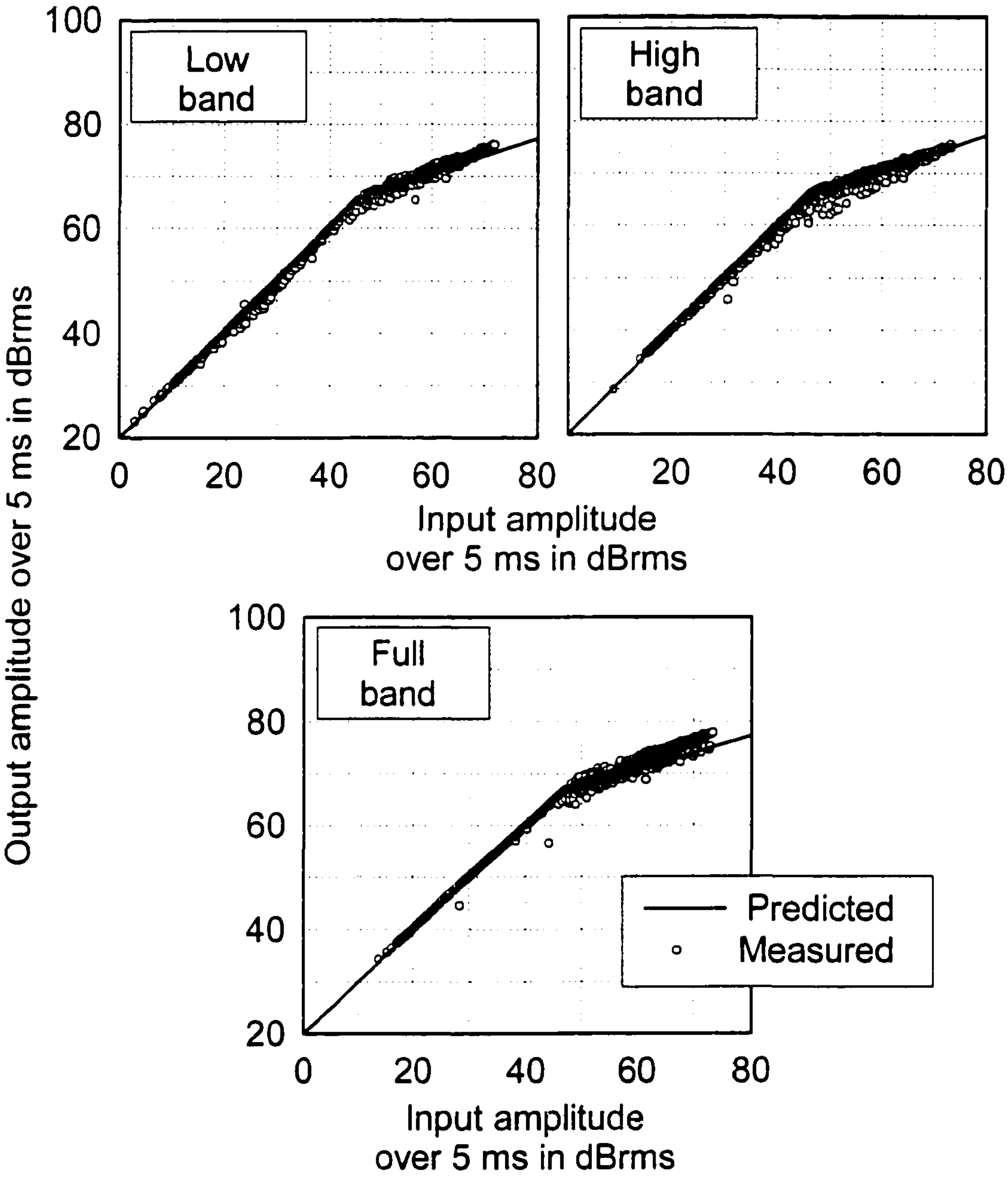

Figure 4.3: Measured and predicted input/output functions for 2-band compression algorithm measured with a 5 ms integration time

Inspection of Figures 4.3 and Table 4.1 reveals that the obtained input/output functions for the LF and HF bands are in close agreement with that predicted. The mean and standard deviation of discrepancy for the LF band were 0.13 dB and 0.68 dB and those for the HF band were -0.21 dB and 0.70 dB. When measured using the full band, the mean and standard deviation increased slightly to 0.85 dB and 1.05 dB, respectively. The data suggest that the fast-release 2-band compression scheme was effective when measured separately in the LF and HF band and that the effectiveness is only marginally reduced when measured using the full band.

Comparison of the fast-release single-band compression and the 2-band compression schemes reveal that the 2-band compression scheme is largely more effective when measured separately in the LF and HF band than the single-band compression scheme. However, when measured using full band, the single-band compression scheme is slightly more effective than the 2-band compression scheme. These findings points out that the desired effectiveness of compression is obtained when the bandwidth and the integration time of the acoustical measure correspond to that used by the compression scheme. The results also point out that the fast-release 2-band compression scheme is more effective than fast-release single-band compression in reducing simultaneous intensity variations that occur across the 2 bands.

2) Effectiveness of fast-release 2-band compression for 125 ms temporal integration time:

Figure 4.4 shows the dynamic input/output functions for the fast-release 2-band compression scheme obtained using a 125 ms temporal integration time. Table 4.1. shows the mean and standard deviation of the discrepancy between the obtained and predicted functions.

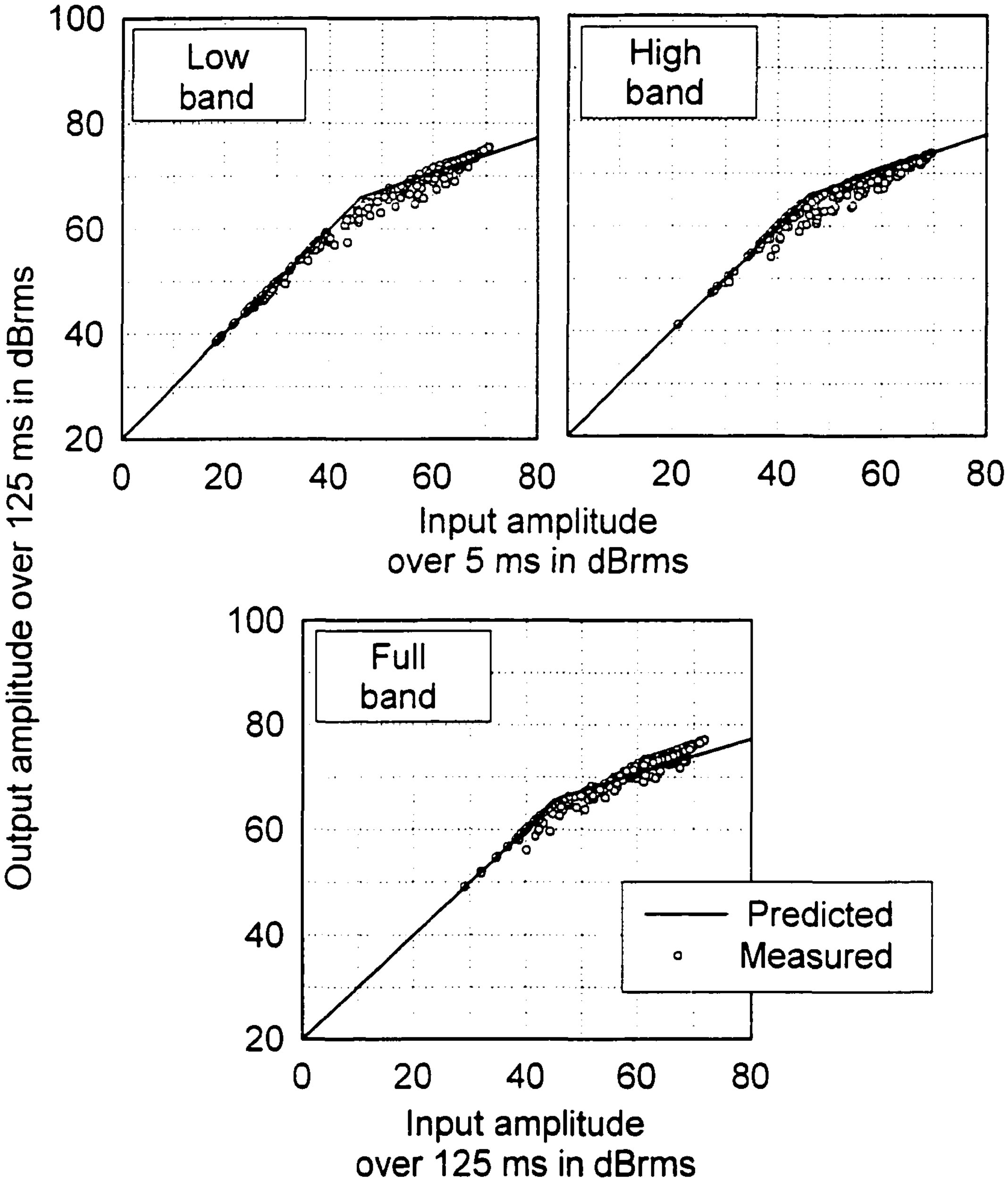

Figure 4.4: Measured and predicted input/output functions for 2-band compression algorithm measured with a 125 ms integration time

Inspection of Figure 4.4 and Table 4.1 reveals that the fast-release 2-band compression scheme is slightly more effective for the LF band followed by full band and HF band. Comparison of dynamic input/output functions obtained using 5 ms and 125 ms temporal integration times reveal no clear-cut variations of the discrepancy as the integration time is increased. This finding is at odds with that noted for fast-release single-band compression scheme, where it was found that the effectiveness of compression decreased with increase in the integration time of the acoustical measure. Since for the fast-release 2-band compression, with the 125 ms integration time, the mean discrepancy for the worst scenario (HF band) was only about 1 dB, and since this value is much lower than that obtained with fast-release single-band compression, it can be reasonably concluded that the effectiveness of the fast-release 2-band compression was not affected with increase in the integration time of the acoustical measure.

4.6. SUMMARY OF FINDINGS FOR COMPRESSION CONDITIONS:

The results of the acoustical analysis of the compression conditions are summarized in Figure 4.5. This figure shows the mean and 95% confidence limits of the dB difference between the obtained and predicted input/output functions for the 2 compression conditions and the 2 temporal integration times.

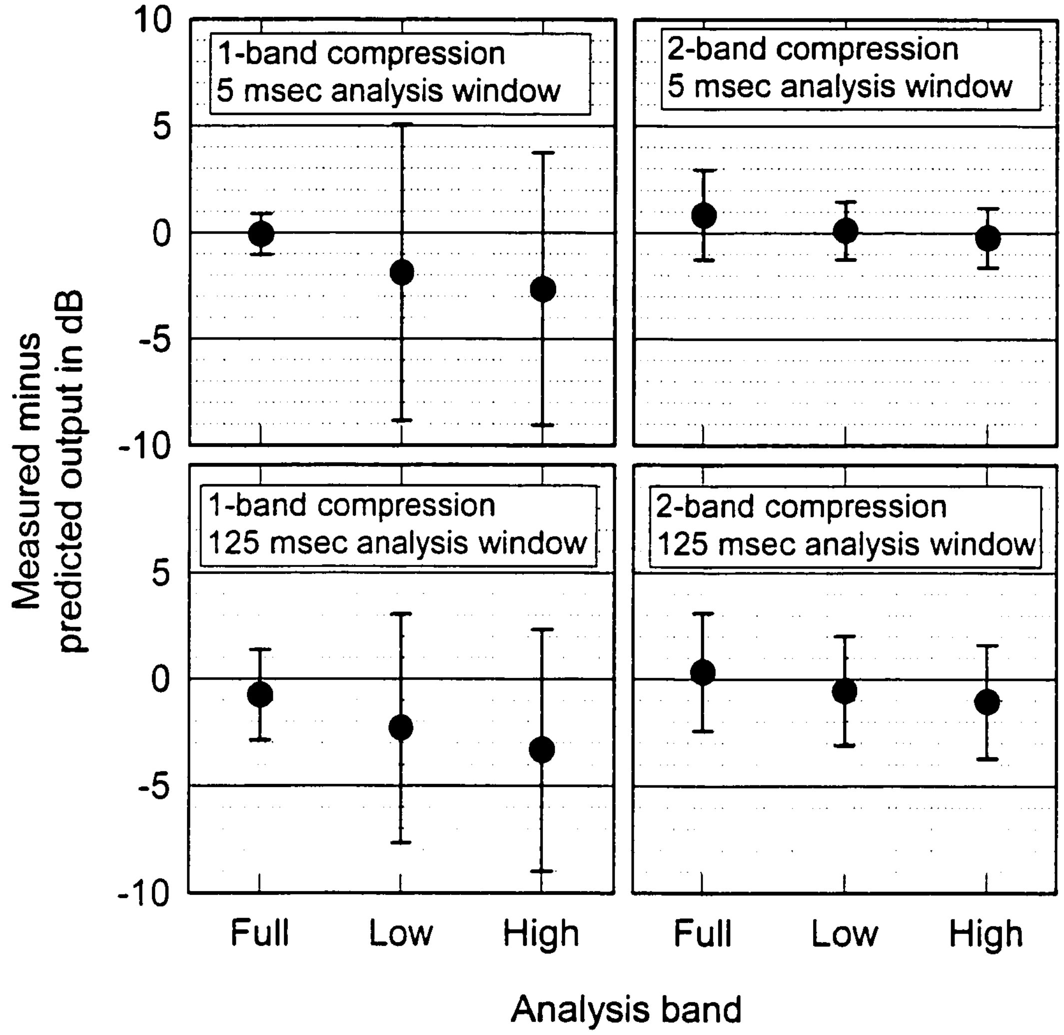

Figure 4.5: Mean and 95% confidence limits of the difference between the obtained and predicted input/output functions.

Inspection of this figure reveals 3 important findings. First, the obtained input/output function is in close agreement with that predicted, provided the temporal integration time and the bandwidth of the acoustical measure correspond to that used by the compression scheme. For example, for the fast-release single-band compression condition, the best results were obtained for full band analysis incorporating 5 ms temporal integration time. For this condition, 95% of the obtained data fell within +0.90 dB and -1.02 dB of that predicted. For the fast-release 2-band compression condition, the best results were obtained for LF and HF band analysis incorporating 5 ms temporal integration time. For this condition, 95% of the obtained data in the LF band fell within +1.49 dB and -1.23 dB of that predicted. In the HF band, 95% of the obtained data fell within +1.19 dB to -1.61 dB of that predicted.

Second, Figure 4.5 reveals that as the integration time of the acoustical measure is increased, there is a general decrease in the effectiveness of the compression schemes in meeting their intended criteria, even when the acoustical measure's frequency bandwidth matched that used by the compression schemes. For example, for full band analysis for the fast-release single-band compression condition, changing the integration time from 5 ms to 125 ms resulted in increasing the dB range of the 95% confidence limits to +1.37 dB and -2.87 dB. Similar results were also obtained for the fast-release 2-band compression condition.

Third, Figure 4.5 reveals that fast-release 2-band compression is more effective than the fast-release single-band compression in reducing simultaneous intensity variations that often occur in the LF and HF bands even when the temporal integration time corresponds to that used by the human ear. For example, for 125 ms integration time, in the LF band, the 95% confidence limits were +3.06 dB and -7.66 dB for the single-band compression condition and only +2.04 dB and -3.12 dB for the 2-band compression condition. Similarly, in the HF band, the 95% confidence limits were +2.33 dB and -8.99 dB for the single-band compression condition and only +1.06 dB and -3.72 dB for the 2-band compression condition.

4.7. DISCUSSION:

The assortment of seemingly contradictory results can mostly be explained. First, for the HF-emphasis-alone condition, although the 8 dB HF-emphasis indeed equalized the 90th percentile cumulative levels of the LF and HF bands within 1 dB, the full band cumulative amplitude distribution does not seem to reflect any significant decrease in the dynamic intensity range of speech. Inspection of the relation of the cumulative amplitude distributions of the LF and HF bands, indeed suggest that the HF-emphasis-alone condition would result in audibility of any given percentile distribution of the HF band at 8 dB lower sensation level than no-processing condition. Therefore, it can be argued that the full band cumulative amplitude distribution is simply not a good reflector of importance functions for speech recognition in the intensity domain.

For the fast-release single-band compression condition, the excellent agreement between the obtained and predicted full band RMS levels when using 5 ms integration time simply confirms that the algorithm was performing as intended. The greater variability when using a 125 ms integration time measurement can be attributed to errors introduced by averaging time periods during which varying amounts of compression have occurred.

The poorer performance of the fast-release single-band compression when assessed in the LF and HF bands separately is also predictable. When the spectrum peaks in the LF band, signal level is changed, not only in the LF band, but also in the HF band. Similarly, when the spectrum peaks in the HF band, level is changed in the LF as well as HF bands. Thus, the processed level in each band depends, sometimes, on energy in the other band - invalidating the predictive formula. In fact, one would predict that the fast-release 2-band compression, in which the adjustment in each band depends only on the level in that band - would lead to greater agreement between the predicted and obtained levels - as was the case.

4.8: CONCLUSION:

These findings confirm that a fast-release 2-band compression scheme is more effective in terms of reducing the dynamic range in the 2 bands than a fast-release single-band compression scheme. They also show increased variability of predictive accuracy when performance is measured with a temporal window similar to the integration time of the ear, rather than the much shorter window used in processing.

In quantitative terms, using 95% confidence limits, the measured output, over 5 ms, in the LF band was within +1.49 dB* and -1.23 dB of predicted for the fast-release 2-band compression and within +5.09 dB and -8.83 dB for the fast-release single-band compression. Measuring over 125 ms, these values change to +2.04 dB and -3.12 dB for the fast-release 2-band compression and +3.06 dB and -7.66 dB for the fast-release single-band compression.

In the HF band, measuring over 5 ms, the output was within +1.19 dB and -1.61 dB for the fast-release 2-band compression and within +3.75 dB and -9.05 dB for the fast-release single-band compression. Measuring over 125 ms, these values change to +1.06 dB and -3.72 dB for the fast-release 2-band compression and +2.33 dB and -8.99 dB for the fast-release single-band compression.

Considering that the goal is to compress the top 30 dB of the signal into a 10 dB range, errors of 7 to 9 dB for the fast-release single-band compression would appear to be serious.

On the basis of these data, one would predict that the fast-release 2-band compression would be more effective than fast-release single-band compression in terms of reducing the effective dynamic range of speech when evaluated behaviorally.

---

* each value is the mean plus or minus 2 standard deviations, using data from table 4.1.

CHAPTER V: PERCEPTUAL VALIDATION

5.1. PURPOSE:

To compare the relative effectiveness of HF-emphasis, fast-release single-band compression, and fast-release 2-band compression in reducing the perceptual short-term dynamic intensity range of speech.

5.2. OUTLINE:

Performance-Intensity (PI) functions for no-processing, HF-emphasis-alone, pre-compression HF-emphasis plus fast-release single-band compression, and pre-compression HF-emphasis plus fast-release 2-band compression conditions were obtained in three groups of 8 pink-noise masked, normal-hearing young adults. The PI functions for the above 4 conditions were obtained in 3 separate experiments[*]. In the first experiment, PI functions for no-processing and HF-emphasis-alone conditions were obtained. In the second experiment, PI functions for HF-emphasis-alone and pre-compression HF-emphasis plus fast-release single-band compression conditions were obtained. In the third and final experiment, PI functions for pre-compression HF-emphasis plus fast-release single-band compression and pre-compression HF-emphasis plus fast-release 2-band compression conditions were obtained. Note that, in order to make reasonable comparisons between the 3 experiments, the

---

[*] Because of the limited availability of test material and subject time, the perceptual validation of the 3 processing conditions were obtained in 3 separate experiments. Each experiment involved different groups of subjects, but the same basic procedure. There were, however, differences in the sensation levels used for testing.

experiments were designed to carry a common condition between the experiments. Specifically, the common condition between experiment 1 and 2 was the HF-emphasis-alone condition, and that between experiment 2 and 3 was the fast-release single-band compression condition. It was reasoned that if the common conditions between the experiments were equivalent in terms of performance then comparisons between uncommon conditions could be made with reasonable validity.

5.3. SPECIFIC QUESTIONS:

1) Does a simple HF-emphasis-alone reduce the perceptual dynamic intensity range of speech?
2) Does the addition of fast-release single-band compression, with a compression ratio of 3:1, reduce the perceptual dynamic intensity range found in the HF-emphasis-alone condition by a factor of 3:1?
3) Does the addition of fast-release 2-band compression, with a compression ratio of 3:1 in both bands, result in greater reduction of the perceptual dynamic intensity range of speech than fast-release single-band compression?
4) Are the measures of the perceptual dynamic intensity range of speech of the common conditions between experiments 1 and 2 and between experiments 2 and 3 comparable?

To answer these questions, perceptual dynamic intensity range was assessed as the intensity range between pairs of correct percent recognition points on the PI function. For each measure the value after processing was compared with that before processing to give an estimate of the compression ratio that could be compared with the target of 3:1.

5.4. METHODS:

5.4.1. Subjects:

A total of 24 adults participated in this study. These subjects were recruited through advertisements in the community. Most of the subjects received monetary compensation for their participation in the study. None of the subjects had any reported complaints of hearing or speech problems. Among the 24 subjects, there were 11 males and 13 females. These 24 subjects were divided into 3 groups of 8 subjects each, and assigned to the 3 experiments. For the first experiment, there were 3 males and 5 females. For the other two experiments, there were 4 males and 4 females. Informed consent was obtained from each subject at the onset of their participation (see Appendix E for details).

5.4.2. Hearing testing:

Each subject underwent a basic pure tone audiometric testing. A 2-channel clinical diagnostic audiometer with TDH-50 headphones mounted on MX-41/AR supraaural cushions were used for testing (GSI-10 from Grason-Stadler, Inc.). Electroacoustic calibration of the audiometer was conducted and found to be in conformity with ANSI (1996) standards for output accuracy and attenuator linearity (see Appendix F).

Thresholds at octave and inter-octave frequencies in the range from 250 to 8000 Hz were obtained using a modified Hughson and Westlake procedure suggested by Carhart and Jerger (1959). In this procedure, testing began at a reasonably loud level and thresholds were bracketed using a 10 dB down and 5 dB up procedure. The lowest intensity level that resulted in correct responses in 2/3 consecutive trials was taken as the threshold of audibility.

Individual subjects' pure tone thresholds in quiet (in dB SPL) for the 3 experiments are shown in Tables G.1a, G.2a, and G.3a, respectively, in Appendix G. Auditory thresholds in dB SPL were derived by using HL to SPL conversion factors for 6 cc coupler. The mean and range of thresholds for groups of subjects for the 3 experiments are shown in Figure 5.1.

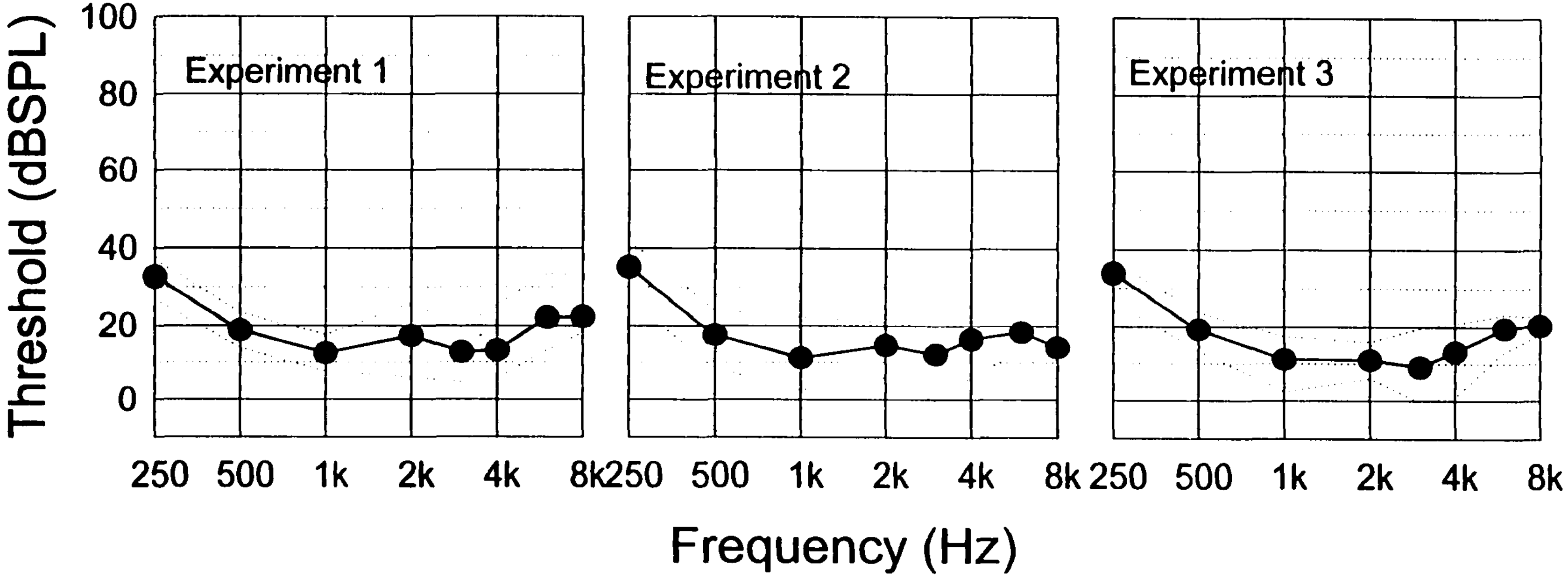

Figure 5.1: Mean and range of pure tone thresholds in quiet for groups of subjects for experiments 1, 2, and 3

As Figure 5.12 shows, the thresholds of audibility vary as a function of frequency. Also, there is a large inter-subject variability in thresholds. Since the thresholds vary as a function of frequency, the sensation levels at the different frequencies will also vary when listening to broadband signals, such as speech. Further, since the short-term speech spectrum varies considerably over time, it is very difficult to predict the interaction between the threshold of audibility curve and the speech spectrum.

The measurement of the perceptual dynamic range of speech using PI functions would require that such listening biases be controlled. A simple way to control such listening biases would be to simulate a flat configuration of hearing thresholds, at least in the speech frequency range, such that the sensation levels at the different frequencies are determined only by the frequency-intensity characteristics of the speech signal and not by individual hearing characteristics

5.4.3. Masking:

In this study, pink noise was chosen as a masker because its 1/3rd octave band level is independent of frequency. The level of the masking noise was chosen to produce, on average, a threshold of 40 dB SPL at all frequencies.

5.4.3.1. Instrument setup for masking:

A pink noise source was obtained by filtering the output of a white noise generator (Coulbourn S81-02) with a pink noise filter (Coulbourn S86-05). This pink noise source was fed to channel-2 of a clinical audiometer as an external input. The dB HL reading of the channel-2 was set to 40 HL. Both pure tones and the pink noise was delivered to the same headphone to obtain ipsilateral pink-noise masked thresholds. The 1/3rd octave band noise levels (A- weighting) at different center frequencies, measured using a sound level meter attached to a 6 c.c. coupler, are shown in Appendix H and Figure 5.2.

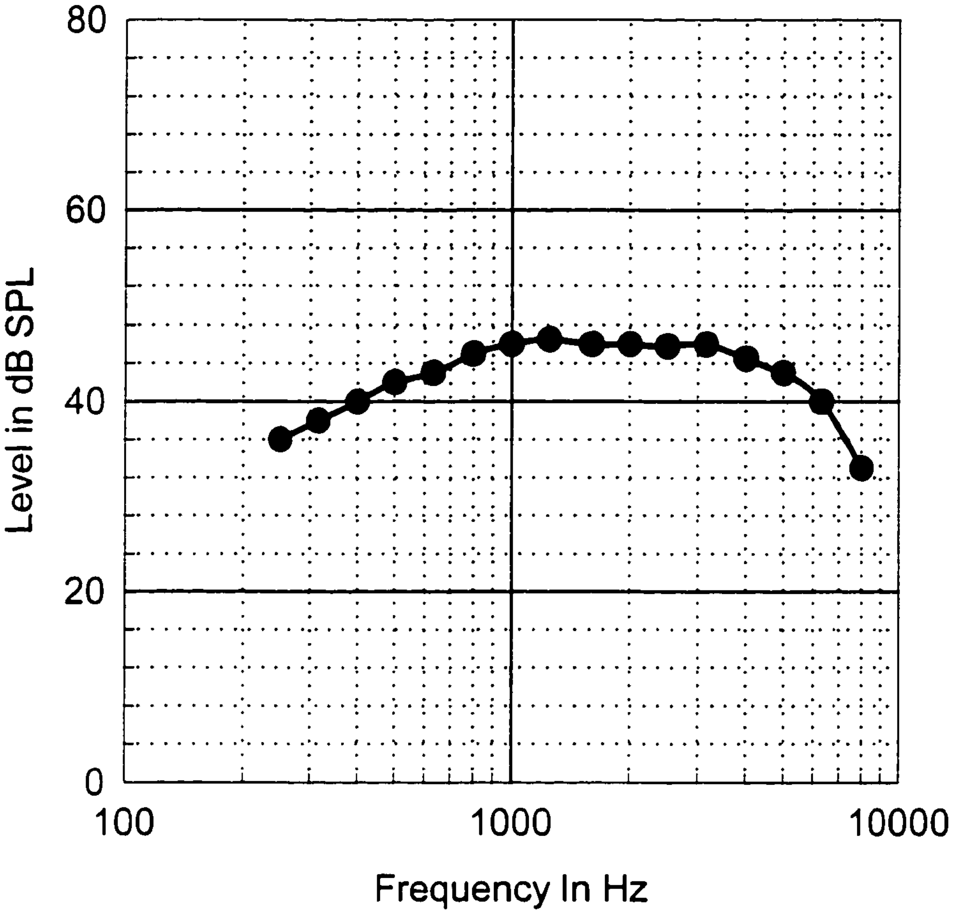

Figure 5.2: One-third octave band spectrum of the pink-noise measured at the output of TDH-50p headphone attached to a 6 cc coupler.

The 1/3rd octave levels were obtained with channel-1 of the audiometer set to speech as the stimulus. When the stimulus setting of the channel-1 was changed to pure tone, variations in the output level of the noise in channel-2 were observed, although its attenuator setting was fixed at 40 dB HL. These variations in the output levels were measured and noted in appendix H. To ensure delivery of exact noise levels during measurement of pure tone masked thresholds as that used for the measurement of the PI function, the needed attenuator setting of the channel-2 (for noise) was determined for each test frequency which resulted in the same noise output as that obtained when speech was chosen for channel-1. Appendix H shows these attenuator settings as well.

5.4.3.2. Masked thresholds:

Pink-noise ipsilateral masked pure tone thresholds at octave and inter-octave frequencies in the range from 250 to 8000 Hz were obtained using the same procedure as that used for measurement of thresholds in quiet. Tables G.1b, G.2b, and G.2c in Appendix G show the individual subject pink-noise masked pure tone thresholds as a function of frequency. The group mean and range of masked thresholds for the 3 experiments are shown in Figure 5.3.

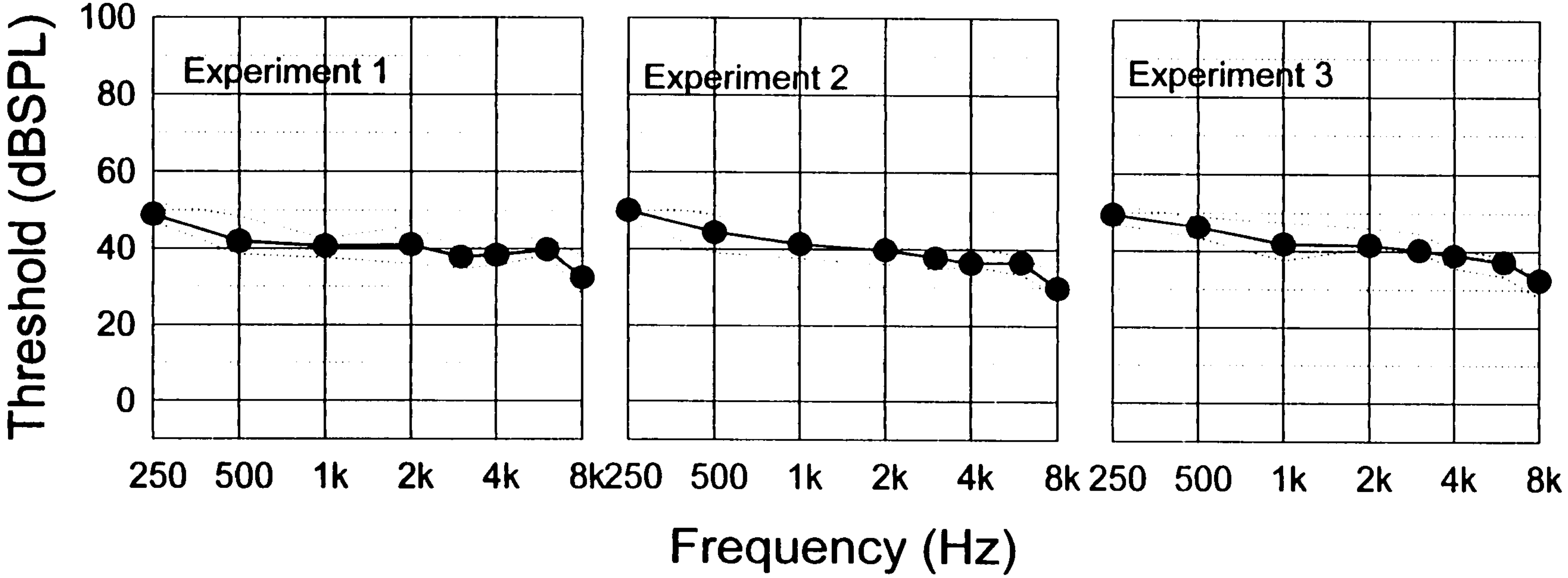

Figure 5.3: Mean and range of pure tone thresholds in noise for groups of subjects for experiments 1, 2, and 3

Inspection of Figure 5.3 shows that the pink-noise masker has resulted in a flat threshold contour of 40 dB SPL in the frequency range between 500-6000 Hz. Further, the inter-subject variability in thresholds have been reduced compared to that obtained in quiet.

5.4.4. Test Stimuli:

Test stimuli consisted of 20 lists of iso-phonemic AB word lists, which were digitally processed to obtain stimuli for no-processing, HF-emphasis alone, pre-compression HF-emphasis plus single-band compression, and pre-compression HF-emphasis plus 2-band compression conditions. (See chapter IV for details of the signal processing techniques). These stimuli were presented with a carrier phase “the word is” which was also processed for the different conditions.

5.4.5. Experiments:

The perceptual validation of single-band and 2-band compression schemes were completed in 3 separate experiments.

Experiment 1: In this experiment, PI functions for no-processing and HF-emphasis-alone conditions were obtained to determine the perceptual dynamic range of unprocessed speech, and to determine the extent to which a simple HF-emphasis alone reduces this range.

Experiment 2: In this experiment, PI functions for HF-emphasis-alone and pre-compression HF-emphasis plus fast-release single-band compression conditions were obtained to determine if single-band compression with a compression ratio of 3:1 reduces the perceptual dynamic range compared with HF-emphasis-alone by a factor of 3:1.

Experiment 3: In this experiment, PI functions for pre-compression HF-emphasis plus fast-release single-band compression and pre-compression HF-emphasis plus fast-release 2-band compression conditions were obtained to determine if 2-band compression with a compression ratio of 3:1 in the LF and HF bands results in greater reduction of the perceptual dynamic range of speech than single-band compression.

5.4.6. Instrumentation setup for the experiments:

The test stimuli were played out from an IBM desktop computer equipped with a 16-bit A/D board. A Windows-based software program using Multimedia Toolbook was written to play the stimuli. This program allowed specification of the condition and list number to be played. It then randomized the list, and saved it as a text file under the subject's name specified. The program permitted the playing of one stimulus word with its corresponding processed carrier phase at a time.

The line output of the A/D board was connected to the tape-A input of channel-1 of a two-channel clinical audiometer (GSI 10). As indicated earlier, a pink noise source was obtained by filtering the output of a white noise generator with a pink noise filter. This pink noise source was fed to channel-2 of the clinical audiometer as an external input. The attenuator of Channel-2 was set at 40 dB HL. The attenuator of channel-1 of the audiometer was used to vary the presentation level of the test stimuli. For each test condition, the appropriate carrier phrase was used to calibrate the VU meter settings. Both the stimuli and the pink noise were delivered to the same headphone for monaural presentation. A 2-channel digital tape recorder was used to record subjects' responses as well as the stimuli.

5.4.7. Presentation levels:

For each experiment, PI functions for 2 conditions were obtained by measuring percent phonemes recognized at 7 different presentation levels. While the 7 presentation levels remained the same for 2 conditions of a given experiment, they did vary across the 3 experiments. The 7 presentation levels for a given experiment were chosen to provide percent phonemes recognized ranging from a guessing score of about 7-10% to a near-perfect score of 90-100%. Further, they where chosen to provide finer resolution of the PI functions at low sensation levels. Therefore, the presentation levels were not ordered on a linear scale. They were bunched together at low presentation levels and relatively more separated at higher presentation levels. At each presentation level, 2 lists of AB words were presented.

The presentation levels, in dB HL, for the 3 experiments were as follows:

Experiment 1: 22, 24, 27, 30, 35, 45, and 55 dB HL

Experiment 2: 20, 23, 26, 30, 34, 38, and 45 dB HL

Experiment 3: 22, 24, 26, 28, 32, 36, and 40 dB HL

5.4.8. Dependent variable:

The dependent variable for the perceptual validation experiments was percent phonemes recognized. Additions, deletions, or substitutions were counted as errors.

5.4.9. Procedure:

Subjects were tested individually under headphones in an IAC, sound attenuating enclosure. All subjects were tested monaurally. The test ear and the presentation order of the 2 conditions were alternated across the 8 subjects.

Test stimuli in a background of pink noise was presented monaurally through a TDH-50p headphone. Test stimuli were presented with a carrier phase “The word is”. Subjects were instructed to repeat and write down what they heard. They were also encouraged to guess if they were not sure of what they heard.

PI functions were obtained using a descending-ascending procedure. Testing began with the presentation of 1 AB word list at the highest chosen presentation level. Following this, the presentation level was decreased 6 times, and at each level 1 AB word list was presented. At the lowest presentation level, 2 lists were presented successively. Then the presentation levels was increased 6 times and at each level 1 AB word list was presented. At the 3 highest presentation levels in the ascending direction, the 3 AB word lists used at the lowest presentation level were re-administered. This procedure ensured that subjects listened to stimuli of the lowest presentation level (poorer performance) later at high presentation level (better performance) and not vice versa. Thus, at each presentation level, 2 lists of 10 words each were used. Also, for a given test condition, 10 lists were used.

5.4.10. Data collection and scoring:

The audio recordings of a given verbal response and their corresponding stimulus were compared to judge the accuracy of the response. They were scored for the correct identification of the stimulus at the phoneme level. When there was ambiguity of verbal response, the written response was compared with the stimulus and used as the final criterion.

5.5. RESULTS:

5.5.1. Experiment 1:

5.5.1.1. Raw data:

The raw data for experiment 1 will be found in Table I.1 in Appendix I. The results are shown in terms of number of phonemes correctly recognized at each presentation level for the 2 conditions: no-processing and HF-emphasis-alone. Data are shown separately for 2 replications.

5.5.1.2. Repeated-measures analysis of variance:

The raw data in Table I.1 were converted into percent correct phoneme recognized scores and then arcsine-transformed to increase homogeneity of variance. The normalized-arcsine transform (NAU) was of the form

y = arcsine(sqrt(x/100))/arcsine(1)*100 .... (1)

Where, y = normalized arcsine transform

x = percent correct score

Note that scores of 0%, 50% and 100% are unchanged by this transform. Repeated-measures analysis of variance was performed on the arcsine transformed data. The results of this analysis are shown in Table 5.1.

Table 5.1: Results of 3-way repeated-measures analysis of variance of arcsine-transformed data of experiment 1 with condition, replication, and presentation level as main factors.

Summary of all effects Condition, Replication, Presentation level						
Source of Variance	Degrees of Freedom	Estimate Mean Effect	Degrees of Freedom Error	Mean Square Error	F Ratio	p - level
Condition	1	4547.38	7	44.83	101.43	<0.0005 **
Replication	1	76.61	7	36.70	2.09	0.1918
Level	6	22551.14	42	54.03	417.39	<0.0005 **
C x R	1	20.25	7	41.49	0.49	0.5074
C x L	6	213.11	42	56.08	3.80	0.0041**
R x L	6	65.42	42	58.02	1.13	0.3632
C x R x L	6	22.59	42	59.47	0.38	0.8877

** $p < 0.01$

The repeated-measures analysis of variance revealed significant main effects of condition and presentation level but not replication. That is, on average, the percent recognized score of the HF-emphasis-alone condition was higher than no-processing condition. Further, the percent phoneme scores differed at various presentation level. However, the average percent phoneme scores did not differ from replication 1 and 2 ($p > 0.05$).

The results also revealed a significant interaction effect between condition and presentation level indicating that the variations of percent phoneme recognized scores with presentation level were significantly different for no-processing and HF-emphasis-alone conditions. However, replication did not have a significant interaction effect with either condition or presentation level ($p > 0.05$). Data are, therefore, collapsed across the 2 replications for all subsequent analysis.

5.5.1.3. Performance-Intensity (PI) Functions:

Mean phoneme recognition is shown as a function of presentation level for the 2 listening conditions in Table 5.2 and illustrated in Figure 5.4.

Table 5.2: Percent phonemes recognized by each subject at each presentation level for no-processing and HF-emphasis-alone conditions. Each data point is an average of 60 phoneme items

		No-Processing							HF-Emphasis-Alone						
		Presentation Levels (dB HL)							Presentation Levels (dB HL)						
Sub	**Ear**	**55**	**45**	**35**	**30**	**27**	**24**	**22**	**55**	**45**	**35**	**30**	**27**	**24**	**22**
S1	R	98.3	88.3	76.7	73.3	30.0	15.0	6.7	100.0	91.7	83.3	70.0	56.7	33.3	20.0
S2	L	98.3	91.7	66.7	50.0	33.3	25.0	13.3	98.3	96.7	91.7	68.3	73.3	35.0	25.0
S3	R	100.0	91.7	76.7	61.7	28.3	28.3	15.0	100.0	96.7	95.0	73.3	65.0	38.3	25.0
S4	L	100.0	95.0	88.3	50.0	36.7	15.0	11.7	100.0	93.3	78.3	73.3	51.7	21.7	21.7
S5	R	98.3	98.3	66.7	45.0	36.7	28.3	11.7	100.0	100.0	83.3	70.0	53.3	31.7	21.7
S6	L	100.0	95.0	90.0	51.7	36.7	20.0	6.7	100.0	96.7	80.0	75.0	60.0	30.0	28.3
S7	R	96.7	90.0	83.3	48.3	30.0	13.3	23.3	96.7	95.0	90.0	73.3	63.3	51.7	15.0
S8	L	100.0	93.3	71.7	48.3	41.7	10.0	10.0	100.0	98.3	86.7	60.0	61.7	38.3	20.0
MEAN		99.0	92.9	77.5	53.5	34.2	19.4	12.3	99.4	96.0	86.0	70.4	60.6	35.0	22.1
SD		1.2	3.2	9.1	9.4	4.5	7.1	5.3	1.2	2.7	5.8	4.8	7.0	8.6	4.1
SE		0.4	1.1	3.2	3.3	1.6	2.5	1.9	0.4	0.9	2.1	1.7	2.5	3.0	1.4

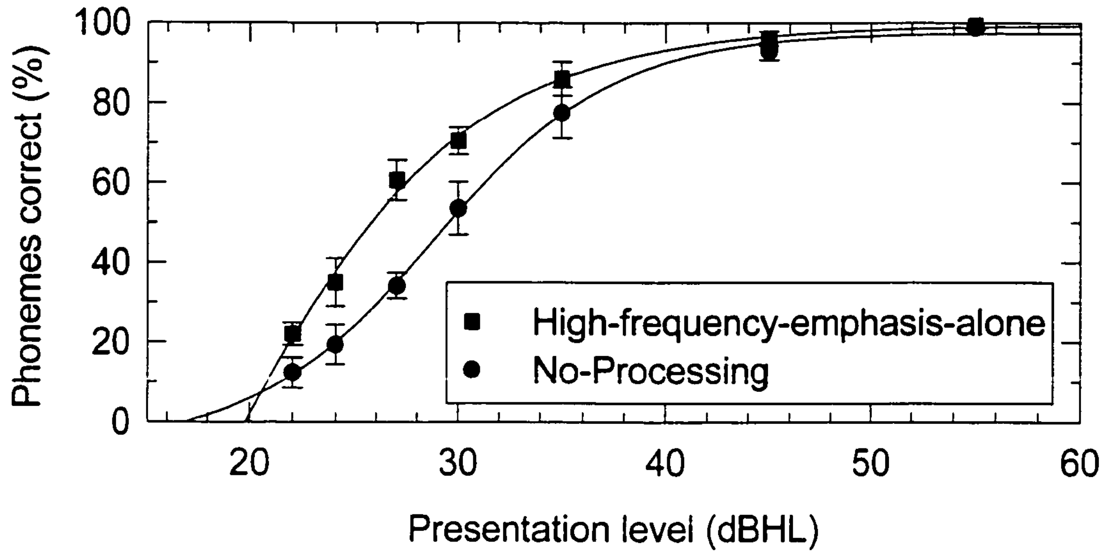

Figure 5.4: Group mean phoneme recognition score (+/-2se) as a function of presentation level for no-processing and high-frequency-emphasis-alone. Curves show least-squares fit to an asymetrical cumulative Gaussian function.

In Figure 5.4, the mean scores were fit to a cumulative Gaussian transition function using a least-squares procedure. The function was:

$y = a + (b-a)/(1+e^{(-(x-c)/d)})$ .... (2)

Where: y = phoneme recognition in %
a = lower asymptote
b = upper asymptote
e = base of natural logarithms
x = presentation level in dB HL (speech)
c = x value for mid-point of the transition function, and
d = a measure of slope

The values of these parameters for the two curves in Figure 5.4 are shown in Table 5.3.

Table 5.3: Values of parameters of the least-squares best-fitting curves of the group mean data of the 2 conditions of experiment 1.

Parameter	Unprocessed	HF-emphasis	Units
a	-6.3	-193.3	%
b	97.5	99.5	%
c	28.8	15.6	dB
d	4.4	6.4	dB^{-1}

It will be seen from Figure 5.4 that the PI function for the no-processing condition has the expected range of about 30 dB in the intensity domain. When measured between the 10% and 90% values, the range is 18.5 dB. The PI function for the HF-emphasis-alone condition also has a range of around 30 dB. When measured between the 10% and 90% values, however, the range is somewhat less than for the no-processing condition - at 16.5 dB. The slopes of the curves, when measured between the 10% and 60% points, are 5.1 and 7.5 %/dB for the no-processing and HF-emphasis-alone conditions, respectively.

The effective compression ratio of the HF-emphasis-alone, measured for different criteria of the perceptual dynamic intensity range of speech, is shown in Figure 5.5. This figure shows the effective compression ratio as a function of the upper intensity limit of the measured dynamic intensity range of speech for different lower limits of the intensity range.

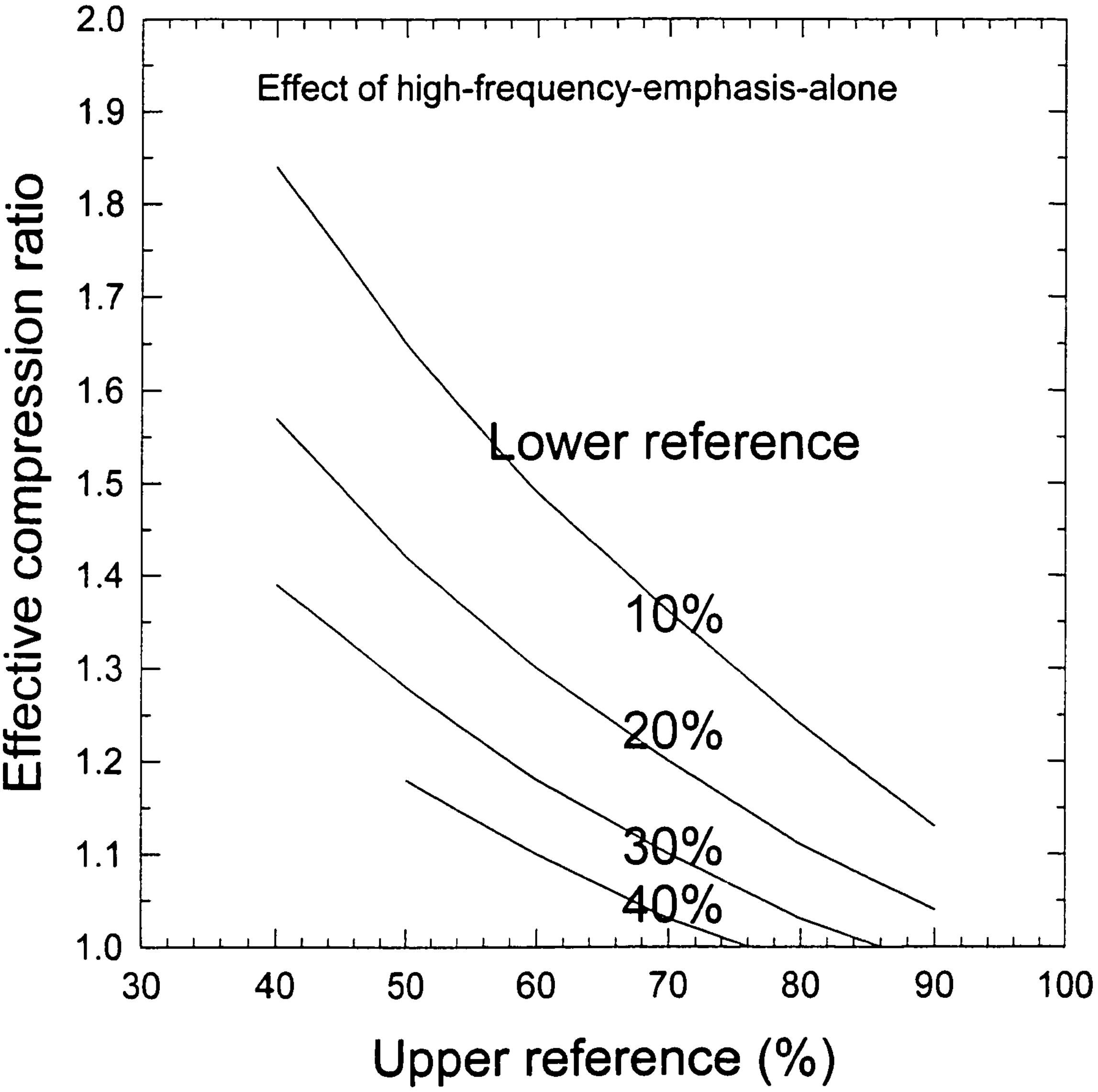

Figure 5.5: Effective compression ratio when introducing high-frequency-emphasis-alone. Ratios are for the dB increase between the lower and upper reference points on the PI functions.

Inspection of Figure 5.5 reveals that the effective compression ratio is reduced as the lower and/or the upper limit of the perceptual dynamic intensity range of speech is increased. This finding suggest that the HF-emphasis-alone is most effective at the lower end of the PI function. Specifically, the highest effective compression ratio is obtained for the perceptual dynamic ranges defined between the 10% and the 60% correct recognition points. Between the 10% and the 60% correct recognition points, the effective compression ratio is about 1.5:1.

Thus, HF-emphasis-alone has the effect of reducing the dynamic range and increasing the slope of the PI function, especially between the 10% and 60% correct recognition points. The major difference between the two sets of data is not so much in terms of the range or slope, however, as in terms of percent scores. Over most of the intensity range, the score with HF-emphasis-alone is some 20 percentage points higher than for no-processing. It will also be seen that the gradual onset at the foot of the PI function for the no-processing condition is missing in the HF-emphasis-alone data. It is not clear, however, whether this is because of a limited lower intensity range of testing for the HF-emphasis-alone condition.

5.5.2. Experiment 2:

5.5.2.1. Raw data:

The raw data for experiment 2 are shown in Table I.2 in Appendix I. The data are shown in terms of number of phonemes correctly recognized at each presentation level for each of the 2 replications for the 2 conditions: HF-emphasis-alone and pre-compression HF-emphasis plus fast-release single-band compression.

5.5.2.2. Repeated-measures analysis of variance:

The data in Table I.2 were converted into percent correct phonemes recognized and transformed into normalized arcsine units. These data were subjected to a repeated-measures analysis of variance. The result of this analysis is shown in Table 5.4.

Table 5.4: Results of 3-way repeated-measures analysis of variance of arcsine-transformed data of experiment 2 with condition, replication, and presentation level as main factors.

Summary of all effects Condition, Replication, Presentation level						
Source of Variance	Degrees of Freedom	Estimate Mean Effect	Degrees of Freedom Error	Mean Square Error	F Ratio	p - level
Condition	1	408.85	7	40.31	10.14	0.0154*
Replication	1	569.2	7	43.00	13.24	0.0083**
Level	6	24209.85	42	51.10	473.82	< 0.0005**
C x R	1	13.96	7	9.77	1.43	0.2710
C x L	6	219.58	42	52.73	4.16	0.0023**
R x L	6	35.12	42	53.03	0.66	0.6803
C x R x L	6	56.9	42	49.24	1.16	0.3481

* $p < 0.05$

** $p < 0.01$

As Table 5.4 shows, there were significant main effects of condition, replication, and presentation level, and significant interaction effect between condition and presentation level. The most important finding of this analysis is the significant main effect of replication. The mean recognition score (collapsed over the 2 conditions) of replication 1 was about 3% higher than replication 2. Note, however, that there was no significant interaction between replication and condition or replication and presentation level. Further, there was no significant 3-way interaction between replication, condition, and presentation level. Therefore, the raw data of this experiment was also collapsed over the 2 replications for all subsequent analysis.

5.5.2.3. Performance-Intensity Functions:

Mean phoneme recognition as a function of presentation level for the 2 listening conditions is shown in Table 5.5 and depicted in Figure 5.6. The mean scores were fitted to a cumulative Gaussian transition function using a least-squares procedure, as in experiment 1. The values of the parameters of this function for the two curves in Figure 5.6 are shown in Table 5.6.

Table 5.5: Percent phonemes recognized by each subject at each presentation level for HF-emphasis-alone and pre-compression HF-emphasis plus single-band compression conditions. Each data point is an average of 60 phoneme items

		HF-Emphasis-Alone							HF-Emphasis + Single-Band Compression						
		Presentation Levels (dB HL)							Presentation Levels (dB HL)						
Sub	**Ear**	**45**	**38**	**34**	**30**	**26**	**23**	**20**	**45**	**38**	**34**	**30**	**26**	**23**	**20**
S1	R	93.3	96.7	85.0	71.7	56.7	26.7	15.0	100.0	98.3	91.7	88.3	65.0	28.3	23.3
S2	L	96.7	86.7	91.7	66.7	46.7	25.0	16.7	96.7	96.7	85.0	80.0	58.3	23.3	6.7
S3	R	91.7	90.0	78.3	65.0	50.0	28.3	13.3	95.0	96.7	90.0	85.0	55.0	25.0	10.0
S4	L	96.7	98.3	81.7	71.7	56.7	36.7	10.0	100.0	90.0	96.7	75.0	56.7	23.3	8.3
S5	R	98.3	95.0	83.3	75.0	41.7	30.0	6.7	100.0	96.7	88.3	81.7	71.7	25.0	3.3
S6	L	96.7	88.3	85.0	78.3	56.7	38.3	8.3	100.0	96.7	83.3	76.7	61.7	26.7	5.0
S7	R	100.0	81.7	88.3	76.7	65.0	33.3	10.0	100.0	98.3	95.0	86.7	60.0	25.0	1.7
S8	L	100.0	86.7	80.0	70.0	43.3	25.0	11.7	96.7	95.0	80.0	81.7	41.7	21.7	5.0
MEAN		96.7	90.4	84.2	71.9	52.1	30.4	11.5	98.5	96.0	88.8	81.9	58.8	24.8	7.9
SD		3.0	5.8	4.4	4.7	8.0	5.2	3.4	2.1	2.7	5.8	4.7	8.7	2.1	6.8
SE		1.0	2.0	1.5	1.6	2.8	1.8	1.2	0.7	0.9	2.0	1.6	3.1	0.7	2.4

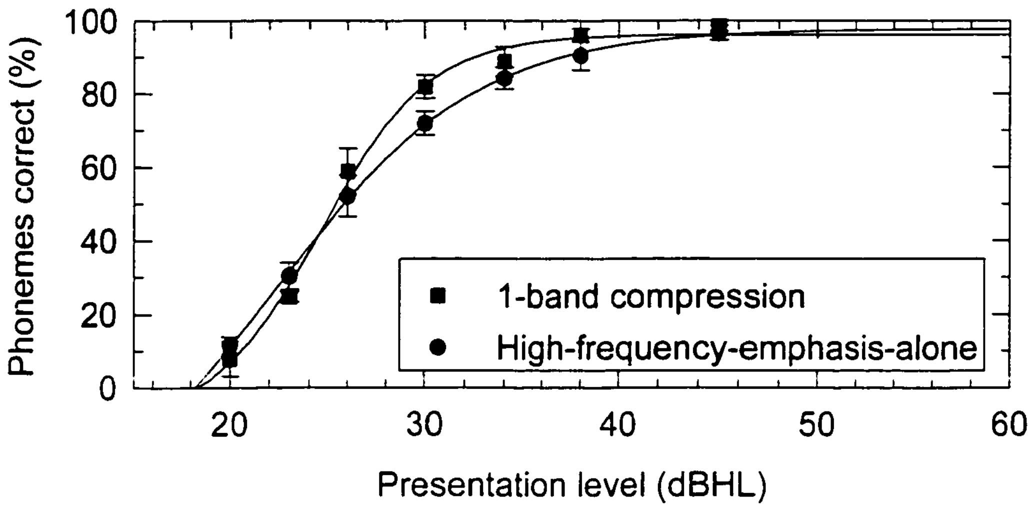

Figure 5.6: Group mean phoneme recognition score (+/-2se) as a function of presentation level for high-frequency-emphasis-alone and 1-band compression. Curves show least-squares fit to an asymetrical cumulative Gaussian function.

Table 5.6: Values of parameters of the least-squares best-fitting curves of the group mean data of the 2 conditions of experiment 2.

Parameter	HF-emphasis-alone	Single-Band compression	Units
a	-45.02	-10.5	%
b	97.8	96.2	%
c	22.3	24.6	dB
d	5.2	2.8	dB^{-1}

The PI function for the HF-emphasis-alone condition shows a range of about 30 dB. The intensity range between the 10% and 90% correct recognition points is about 17.2 dB, which is in close agreement with the 16.5 dB value that was obtained for this condition in experiment 1. In contrast, the PI function for the fast-release single-band compression shows a range of only about 18 dB. When measured between the 10% and 90% correct recognition points, the range is about 11.9 dB. The slopes of the PI functions, when measured between the 10% to 90% correct recognition points, are 4.7 and 6.8 %/dB for the HF-emphasis and single-band compression conditions, respectively.

The effective compression ratio of the single-band compression for different criteria of the perceptual dynamic intensity range of speech is shown in Figure 5.7.

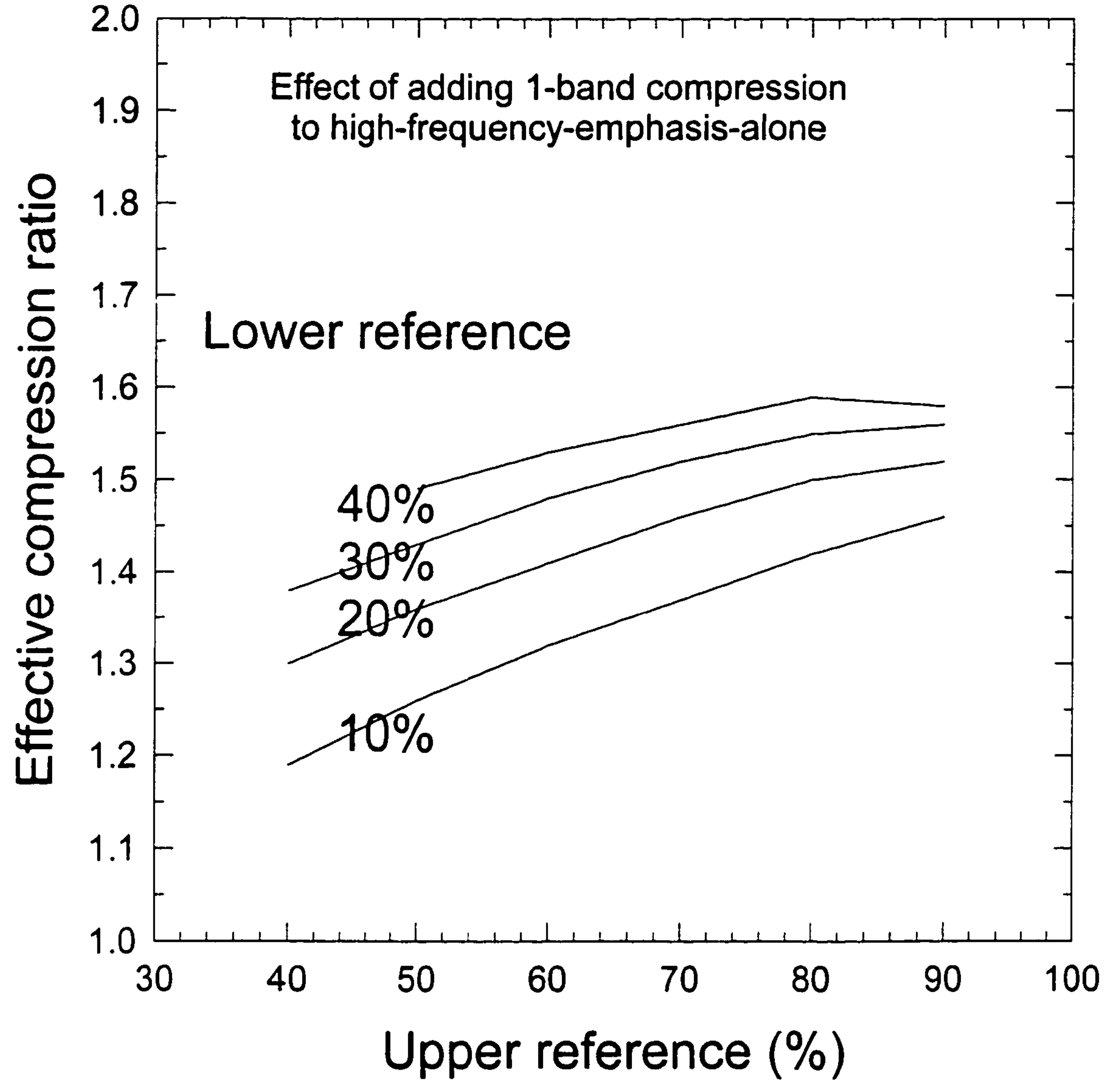

Figure 5.7: Effective compression ratio when changing from high-frequency-emphasis-alone to high-frequency-emphasis plus single-band compression. Ratios are for the dB increase between the lower and upper reference points on the PI functions.

Inspection of Figure 5.7 shows that the effective compression ratio is increased as the lower and/or the upper limit of the perceptual dynamic intensity range is increased. The highest effective compression ratio is obtained for the perceptual dynamic intensity range measured between the 40% and 90% correct recognition points suggesting that the fast-release single-band compression is most effective at the top end of the PI function. Note that this finding is contrary to that obtained when HF-emphasis-alone condition was compared with no-processing condition. Between the 40% and 90% correct recognition points, the effective compression ratio of the single-band compression was 1.6:1

Thus, the fast-release single-band compression has the effect of reducing the overall dynamic intensity range and increasing the slope of the PI function. However, the extent of these effects is less than that intended. The major differences between the two sets of data is the improved performance with fast-release single-band compression, especially between the 40% and 90% correct recognition points. Indeed, post-hoc LSD test for the arcsine transformed data of the repeated-measures analysis of variance revealed that percent recognized scores at 20, 30 and 38 dB HL presentation levels were significantly different between the two conditions. At the 30 and 38 dB HL intensity levels, performance with the fast-release single-band compression was significantly higher than with HF-emphasis-alone, and at the 20 dB HL intensity level, the performance with the HF-emphasis-alone condition was significantly higher than with fast-release single-band compression. The improved performance with HF-emphasis-alone at 20 dB HL level may not necessarily indicate the superiority of HF emphasis, but rather it might have occurred due to errors introduced in defining the actual presentation levels for the two conditions. Note that, as described in section 5.4.6, the presentation levels were calibrated using the carrier phrase of the respective conditions. This presentation level issue would suggest the need to normalize the two sets of data at a low recognition point, e.g. 10% point, for better depiction of the effects of the single-band compression. The effects of this adjustment will be described later in this chapter.

5.5.3. Experiment 3:

5.5.3.1. Raw data:

The raw data for experiment 3 are shown in Table I.3 in Appendix I. The results are shown in terms of number of phonemes correctly recognized at each presentation level for the 2 conditions: pre-compression HF-emphasis plus fast-release single-band compression and pre-compression HF-emphasis plus fast-release 2-band compression. Data are shown separately for 2 replications.

5.5.3.2. Repeated-measures analysis of variance:

The data in Table I.3 were converted into percent phoneme correctly recognized, and transformed into normalized arcsine values. The transformed data were subjected to a repeated-measures analysis of variance. The result of this analysis is shown in Table 5.7.

Table 5.7: Results of 3-way repeated-measured analysis of variance of experiment 3 with condition, replication, and presentation level as main factors.

Summary of all effects Condition, Replication, Presentation level						
Source of Variance	Degrees of Freedom	Estimate Mean Effect	Degrees of Freedom Error	Mean Square Error	F Ratio	p - level
Condition	1	64.24	7	32.90	1.95	0.2050
Replication	1	15.15	7	40.49	0.37	0.5600
Level	6	20562.46	42	45.19	455.03	< 0.0005**
C x R	1	5.36	7	34.49	0.16	0.7051
C x L	6	37.81	42	68.23	0.55	0.7639
R x L	6	15.73	42	63.96	0.25	0.9583
C x R x L	6	33.74	42	39.56	0.85	0.5370

** $p < 0.05$

Inspection of Table 5.7 revealed that neither condition nor replication were significant factors. That is, the average scores obtained with fast-release single-band compression were not significantly different from those obtained with fast-release 2-band compression. Further, the average scores obtained in replication 1 were not significantly different from that of replication 2. As expected, presentation level was a significant factor indicating that the average phoneme scores differed for various presentation levels ($p < 0.0005$). None of the interaction effects were significant. Thus, data for this experiment were collapsed across the 2 replications for all subsequent analysis.

5.5.3.3: Performance-Intensity Functions:

Mean phoneme recognition is shown as a function of presentation level for the 2 listening conditions in Table 5.8 and illustrated in Figure 5.8. The mean scores in Figure 5.8 were fitted with a cumulative Gaussian transition function using a least-squares procedure of the same form as used in experiments 1 and 2. The values of these parameters for the two curves in Figure 5.8 are shown in Table 5.9.

Table 5.8: Percent phonemes recognized by each subject at each presentation level for pre-compression HF-emphasis plus single-band and pre-compression HF- emphasis plus 2-band compression conditions. Each data point is an average of 60 phoneme items

		HF-Emphasis + Single-Band Compression							HF-Emphasis + 2-Band Compression						
		Presentation Levels (dB HL)							Presentation Levels (dB HL)						
Sub	**Ear**	**40**	**36**	**32**	**28**	**26**	**24**	**22**	**40**	**36**	**32**	**28**	**26**	**24**	**22**
S1	R	100.0	91.7	100.0	68.3	58.3	36.7	23.3	95.0	96.7	90.0	80.0	61.7	46.7	16.7
S2	L	96.7	95.0	88.3	56.7	56.7	31.7	16.7	98.3	93.3	86.7	76.7	56.7	38.3	20.0
S3	R	100.0	98.3	86.7	80.0	58.3	45.0	10.0	100.0	96.7	88.3	80.0	50.0	50.0	15.0
S4	L	96.7	91.7	96.7	70.0	43.3	26.7	16.7	100.0	98.3	78.3	66.7	51.7	36.7	15.0
S5	R	96.7	100.0	85.0	71.7	46.7	46.7	1.7	100.0	93.3	95.0	73.3	58.3	43.3	15.0
S6	L	98.3	96.7	95.0	80.0	63.3	50.0	25.0	100.0	98.3	90.0	80.0	56.7	33.3	20.0
S7	R	98.3	90.0	85.0	75.0	53.3	38.3	15.0	96.7	93.3	91.7	75.0	61.7	36.7	13.3
S8	L	98.3	95.0	88.3	70.0	45.0	31.7	15.0	93.3	95.0	90.0	76.7	61.7	35.0	15.0
MEAN		98.1	94.8	90.6	71.5	53.1	38.3	15.4	97.9	95.6	88.8	76.0	57.3	40.0	16.3
SD		1.4	3.5	5.8	7.5	7.3	8.3	7.3	2.6	2.2	4.9	4.5	4.5	6.0	2.5
SE		0.5	1.2	2.0	2.6	2.6	2.9	2.6	0.9	0.8	1.7	1.6	1.6	2.1	0.9

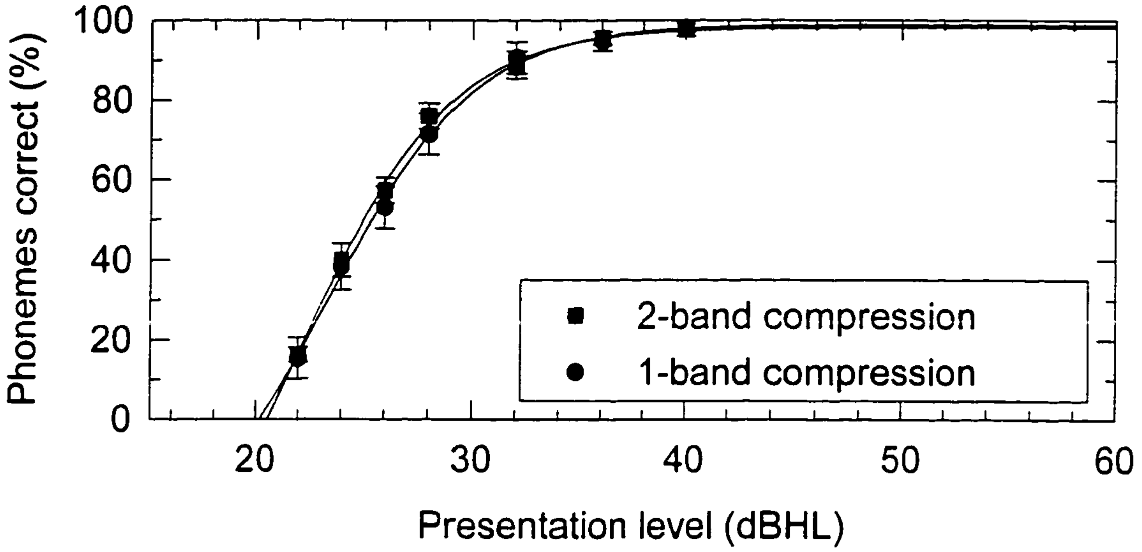

Figure 5.8: Group mean phoneme recognition score (+/-2se) as a function of presentation level for 1-band and 2-band compression.
Curves show least-squares fit to an asymetrical cumulative Gaussian function.

Table 5.9: Values of parameters of the least-squares best fitting-curves of the group mean data of the 2 conditions of experiment 3.

Parameter	Single-Band compression	2-Band compression	Units
a	-35.7	-60.7	%
b	98.9	98.3	%
c	23.6	22.2	dB
d	3.3	3.4	dB^{-1}

Inspection of Figure 5.8 reveals that the PI function for the fast-release single-band compression condition has a intensity range of about 18.5 dB. When measured between the 10% and 90% correct recognition points, the intensity range is about 11 dB and the slope is about 7.3 %/dB. These findings are in good agreement with that obtained for this condition in experiment 2. The fast-release 2-band compression condition also has an intensity range of about 18 dB. Further, the intensity range from the 10% and 90% correct recognition points is about 10.6 dB. Also, the slope between these 2 recognition points is about 7.6 %/dB. Essentially, the PI functions of the 2 listening conditions overlap over the entire intensity range.

The effective compression ratio of fast-release 2-band compression over fast-release single-band compression is shown in Figure 5.9.

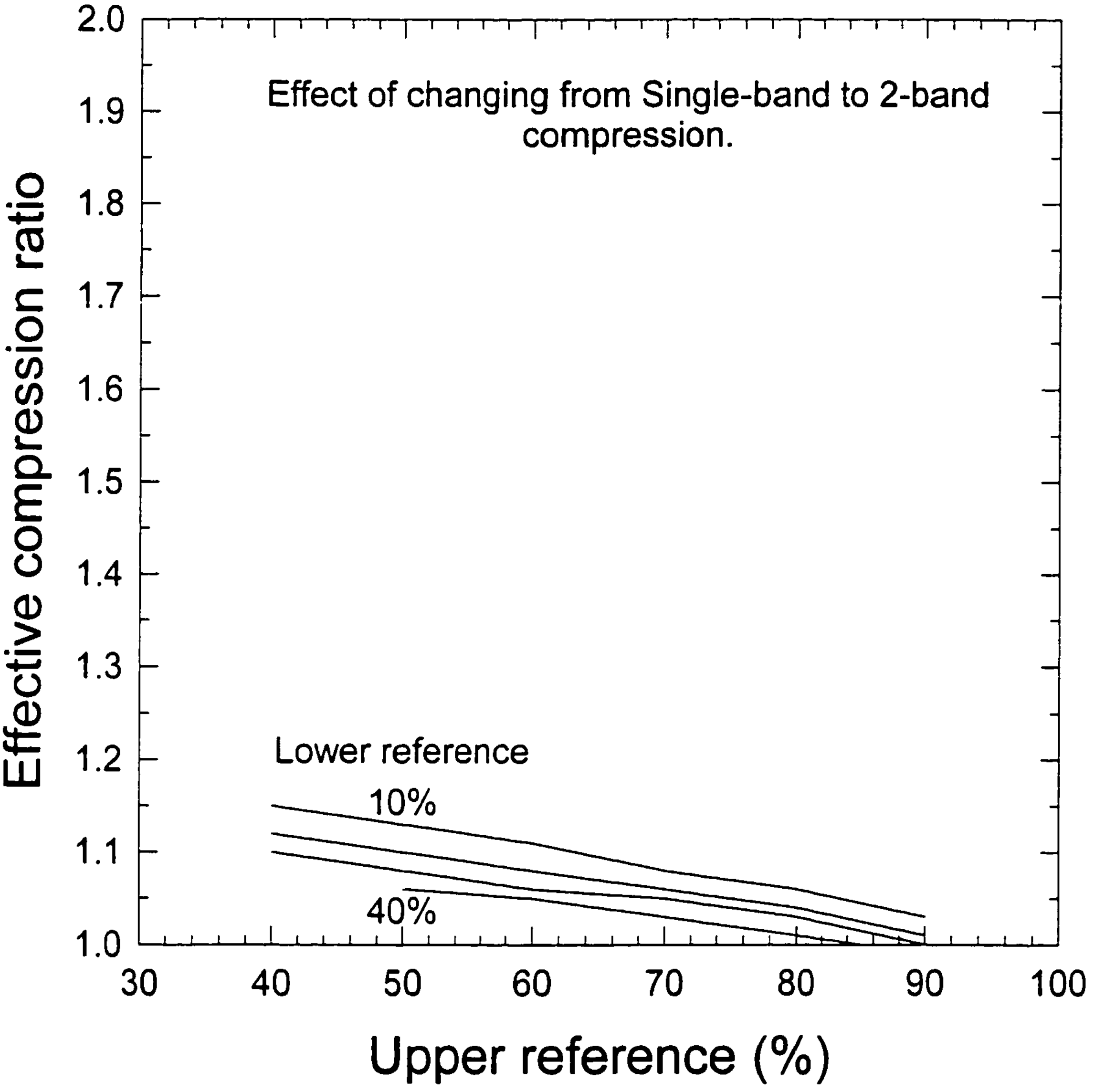

Figure 5.9: Effective compression ratio when changing from single-band to 2-band compression. Ratios are for the dB increase between the lower and upper reference points on the PI functions.

Inspection of Figure 5.9 reveals that there is only a small increase, if any, in the effective compression ratio as we move from single-band compression to 2-band compression. Any advantage of 2-band compression over single-band compression seem to lie at the toe of the PI function. Between the 10% and 60% correct recognition points, the effective compression ratio of the 2-band compression is only about 1.15:1 to 1.10:1. Thus, there are no notable differences between the PI functions for the fast-release single-band and fast-release 2-band compression conditions.

5.5.4. Group Equivalency:

The PI functions of the common conditions between experiment 1 and 2 and between experiment 2 and 3 were compared to provide an estimate of the group equivalency. The common condition between experiment 1 and 2 was the HF-emphasis-alone condition and that between experiment 2 and 3 was pre-compression HF-emphasis plus fast-release single-band compression.

The PI functions for HF-emphasis-alone condition obtained in experiment 1 and 2 are shown in Figure 5.10.

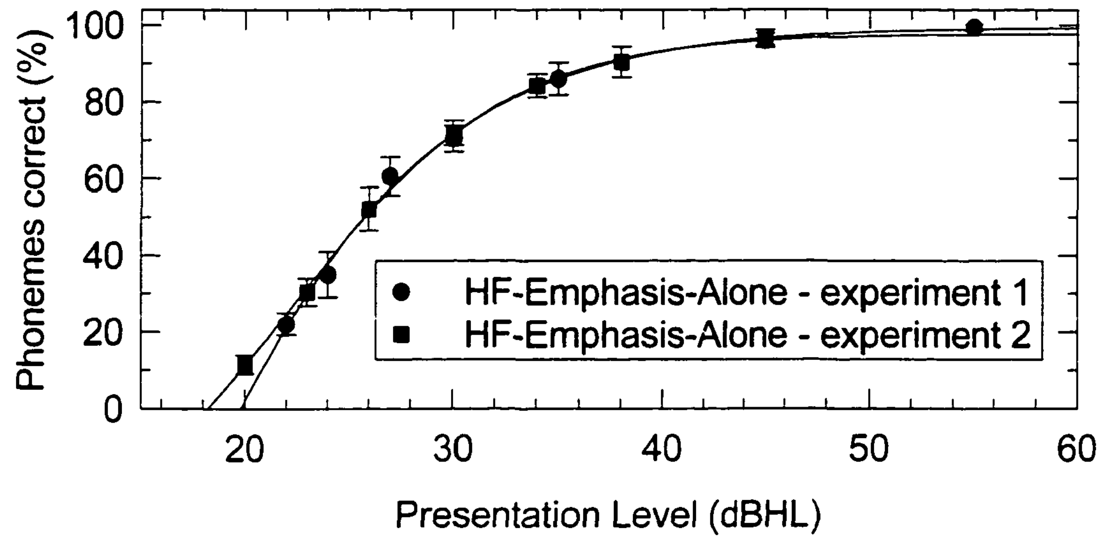

Figure 5.10: Group mean recognition score (+/-2se) as a function of presentation level for HF-emphasis-alone obtained in experiments 1 and 2. Curves show least-squares fit to an asymetrical cumulative gaussian function.

The 2 curves in Figure 5.10 essentially overlap over the entire intensity range. The intensity range between the 10% and 90% correct recognition points were 16.5 dB and 17.2 dB for experiment 1 and 2, respectively. The slope of the curves between the 10% and 60% correct recognition points were 7.5 and 6.5%/dB, respectively. The PI functions for the HF-emphasis-alone condition obtained in experiments 1 and 2 are essentially identical.

The PI functions for fast-release single-band compression condition obtained in experiment 2 and 3 are shown in Figure 5.11.

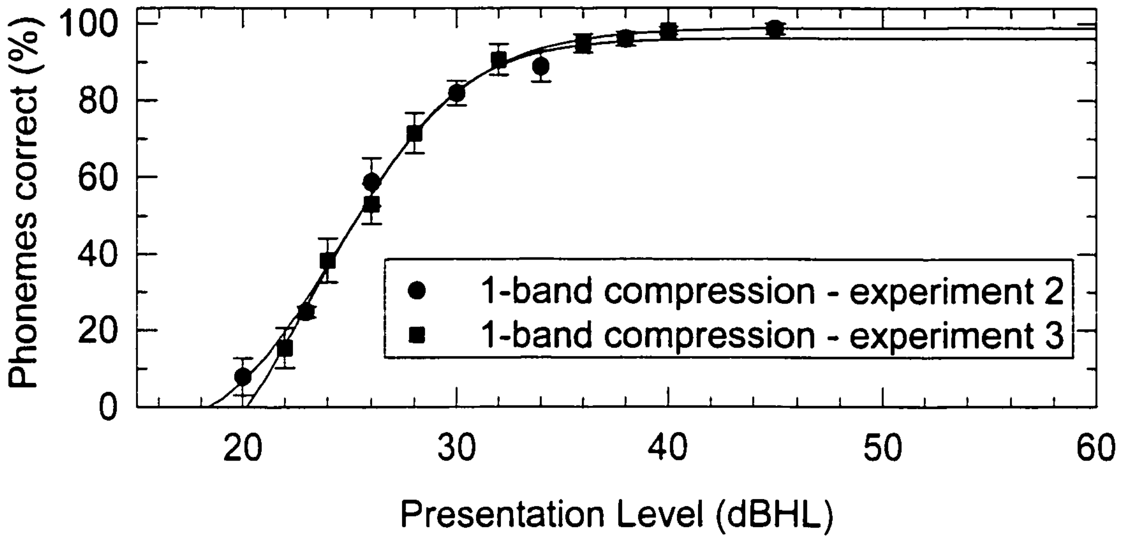

Figure 5.11: Group mean recognition score (+/-2se) as a function of presentation level for single-band compression obtained in experiments 2 and 3. Curves show least-squares fit to an asymetrical cumulative gaussian function.

As with the PI functions for HF-emphasis-alone condition, the 2 curves for single-band compression overlap. The intensity range from the 10% to 90% correct recognition points were 12 and 11 dB for experiment 1 and 2, respectively. The slopes of the curves, between the 10% and 90% correct recognition points, were 6.8 and 7.3%/dB, respectively. These findings suggest that the PI functions for the fast-release single-band compression condition obtained in experiments 2 and 3 are essentially similar.

Since the PI functions of the common conditions between experiment 1 and 2 and between experiments 2 and 3 are similar, it can be argued that the 3 groups of subjects are equivalent. Therefore, comparisons between the uncommon conditions across the 3 experiments can be made with reasonable validity.

5.5.5. Comparisons of uncommon conditions:

Figure 5.12 shows the PI functions for all 4 conditions. In this figure, small intensity adjustments of 1.5 dB or less were made to the functions so that the curves correspond at the 10% correct recognition point. The curves for HF-emphasis-alone and pre-compression HF-emphasis plus fast-release single-band compression conditions were obtained after collapsing the data from experiments 1 and 2, and experiments 2 and 3, respectively.

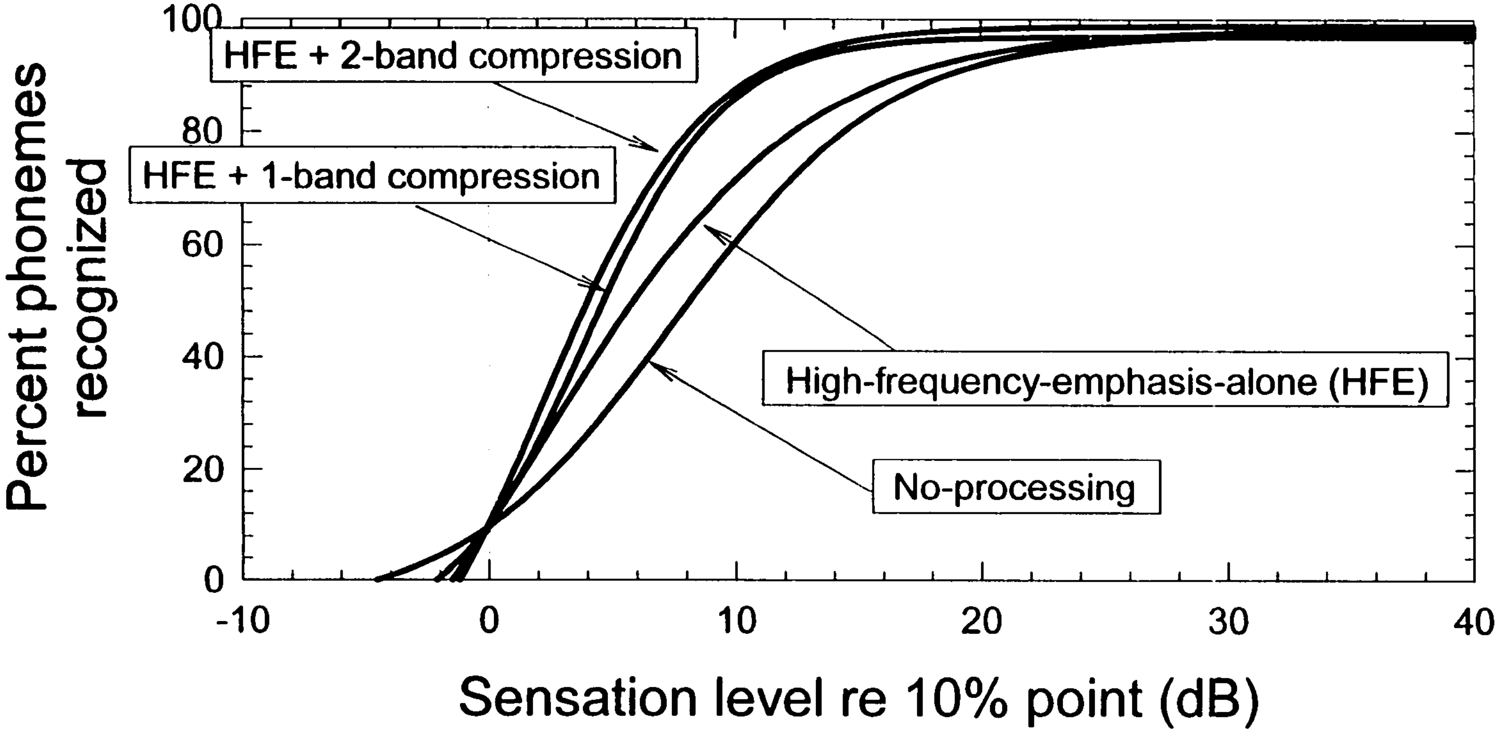

Figure 5.12. Performance/Intensity functions for 4 experimental conditions. Curves are least-squares fits to the the data from 3 experiments. Small adjustments have been made to the dB values so that the functions correspond at the 10% point.

5.5.5.1. No-processing Vs fast-release single-band compression:

The PI function for the no-processing condition shows an overall dynamic range of about 30 dB. Pre-compression HF-emphasis plus fast-release single-band compression has reduced this range to about 18 dB. When measured between the 10% to 90% correct recognition points, the intensity range is about 18.5 dB for the no-processing condition and about 10.9 to 11.9 dB for the fast-release single-band compression. The effective compression ratio of fast-release single-band compression over no-processing, for different perceptual dynamic intensity range of speech, is shown in Figure 5.13.

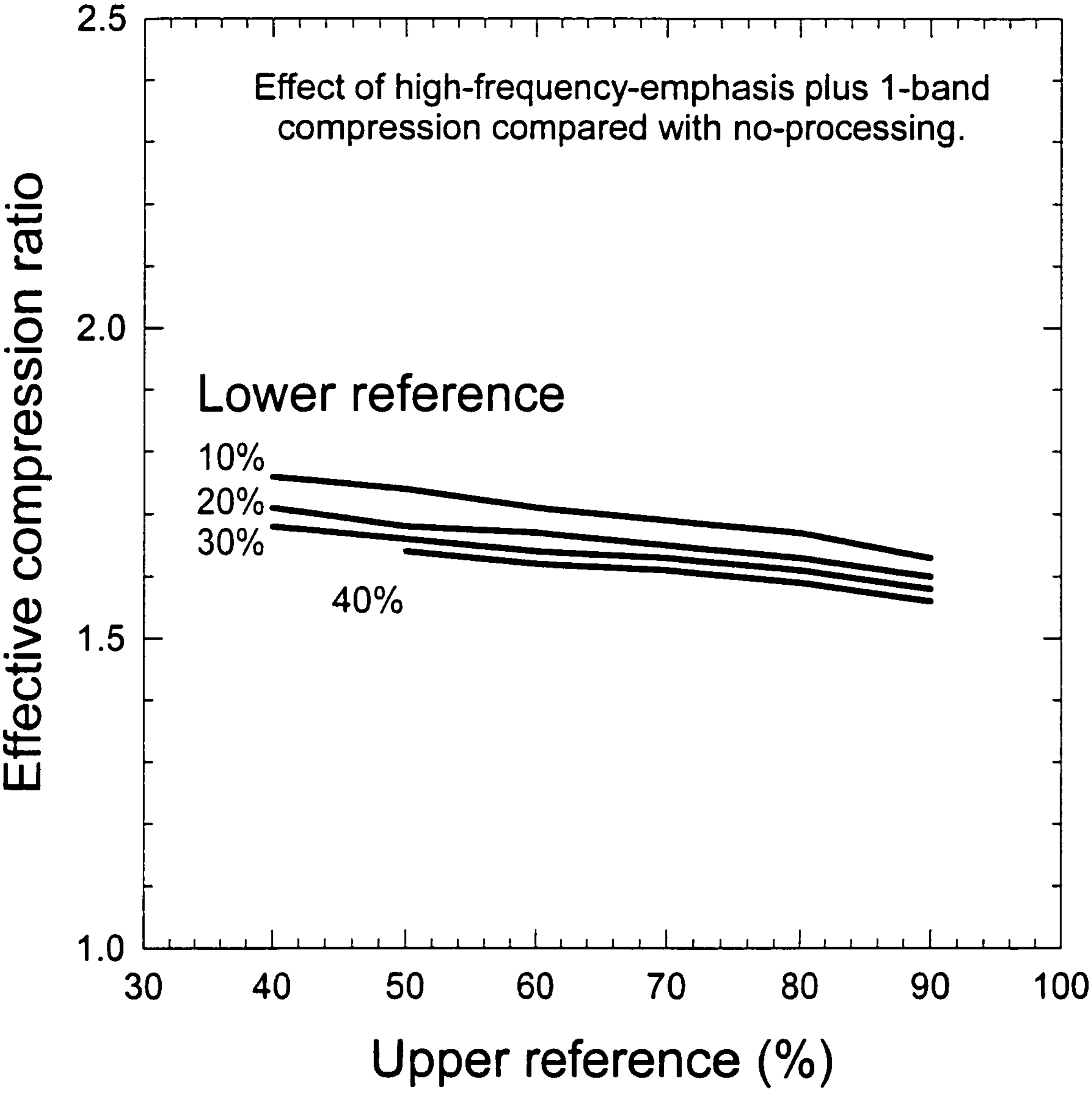

Figure 5.13: Effective compression ratio produced by a combination of high-frequency-emphasis and single-band compression. Ratios are for the dB increase between the lower and upper reference points on the PI functions.

Inspection of Figure 5.13 reveals that the fast-release single-band compression resulted in reducing the dynamic range of the no-processing condition by roughly 1.6 to 1.7 over most of the PI function. The effective compression ratio tends to decrease slightly with increases in the lower and/or the upper limits of the perceptual dynamic intensity range of speech. However, this decrease in effective compression is much lower than that obtained when HF-emphasis-alone condition was compared with no-processing condition.

5.5.5.2. No-processing Vs fast-release 2-band compression:

Comparison of PI functions of no-processing and fast-release 2-band compression conditions reveal that the 30 dB overall dynamic range of the no-processing condition was reduced to about 18 dB by the fast-release 2-band compression. This finding is similar to that obtained when no-processing condition was compared with fast-release single-band compression condition. However, the major advantage of the 2-band compression over single-band compression seem to lie at the toe of the PI function. The effective compression ratio of the fast-release 2-band compression over no-processing condition is shown in Figure 5.14.

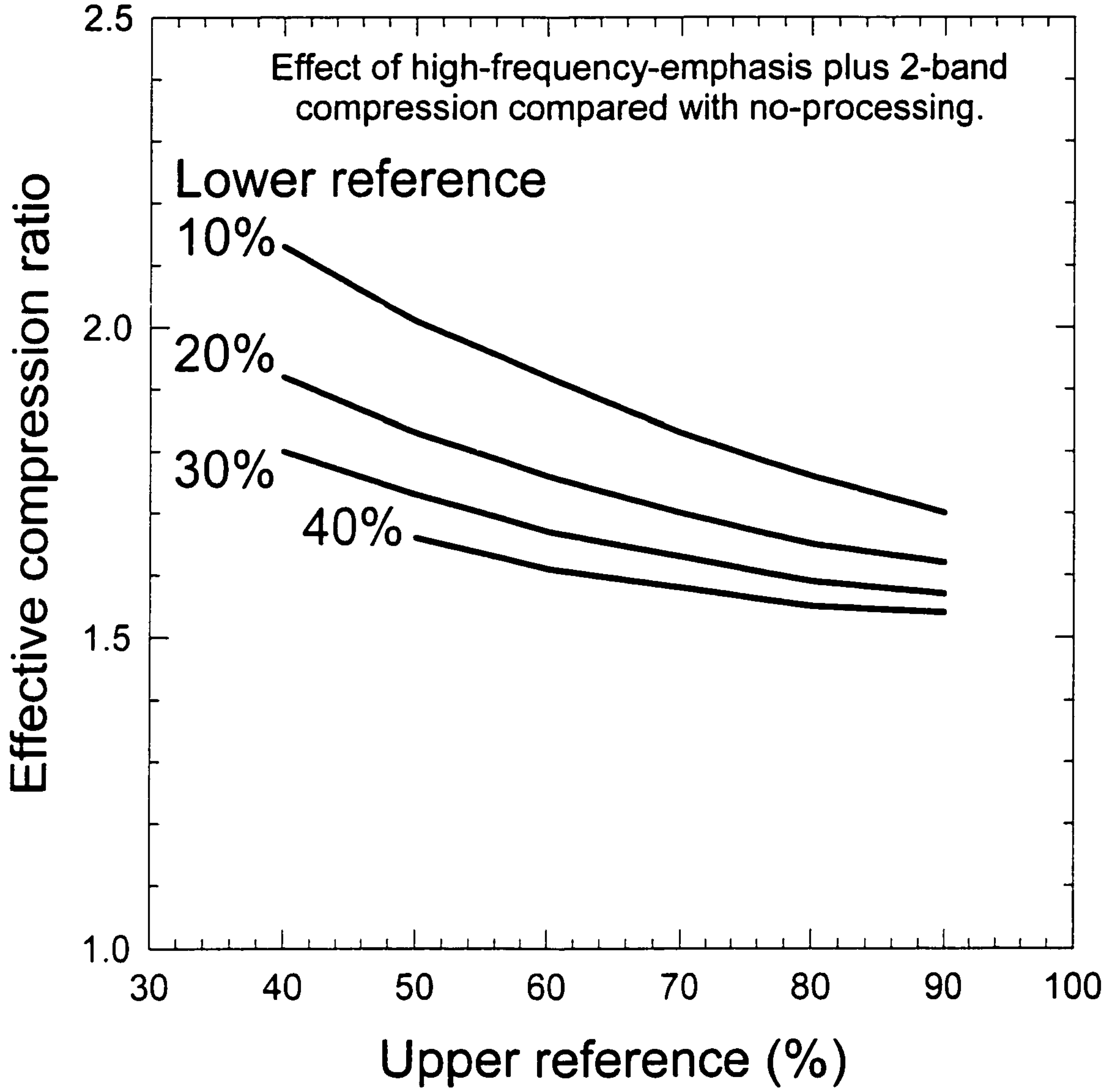

Figure 5.14: Effective compression ratio produced by a combination of high-frequency-emphasis and 2-band compression. Ratios are for the dB increase between the lower and upper reference points on the PI functions.

Inspection of Figure 5.14 reveals that the effective compression ratio was higher especially between the 10% and 60% correct recognition points. Between these 2 recognition points, the effective compression ratio was about 2:1. These results are only slightly superior than those obtained when fast-release single-band compression was compared with no-processing condition. Therefore, these results again reveal that pre-compression HF-emphasis plus fast-release 2-band compression has not produced the expected extent of the reduction of dynamic range of speech compared with that provided by pre-compression HF-emphasis plus fast-release single-band compression.

5.6. DISCUSSION:

The results of the 3 experiments fall mostly in predictable lines. The results of experiment 1 revealed that the 8 dB HF-emphasis has indeed resulted in increasing the slope of the PI function. It was assumed apriori that the dynamic intensity range of speech for speech understanding in the low- and high-frequency bands are roughly 30 dB. Since the peak levels in the low- and high-frequency bands were about 8 dB apart, it was hypothesized that by equating the peak levels of the 2 bands, there will be an overall reduction in the dynamic intensity range of speech perhaps to 22 dB. The results of this study, however, showed that the HF-emphasis-alone did not decrease the overall dynamic intensity range of speech when measured at the 10% and 90% correct recognition points. This finding suggests that the dynamic range of the high-frequency band may be smaller than 30 dB, and that any amplification of the high-frequency band does not decrease the overall dynamic range. Nevertheless, the 8 dB high-frequency boost does increase the perception of speech, especially between the 10% and 60% recognition points, as revealed by the increased slope and an effective compression ratio of 1.5:1. These results indicate that simple HF-emphasis, aimed at equating the peak levels of the low- and high-frequency bands, does not reduce the overall dynamic range of speech. These results indicate that wide dynamic range compression may be necessary for the reduction of the dynamic range of speech, as dictated by the amplification needs of individuals with severe and profound hearing loss individuals.

Addition of fast-release single-band compression, with a compression ratio of 3:1, to pre-compression HF-emphasis had the effect of reducing the overall dynamic range of speech and increasing the slope of the PI function. The overall dynamic range of about 30 dB of the HF-emphasis-alone condition was reduced to about 18 dB by the addition of fast-release single-band compression. However, the effective compression ratio of the fast-release single-band compression condition over HF-emphasis-alone was less than that targeted. Similar lower effective compression ratios for wide dynamic range compression have been previously reported, especially for acoustical measures. This study has demonstrated the effect at both acoustical and perceptual levels. The relation between the acoustical and perceptual measures will be discussed in the next chapter.

PI functions for the pre-compression HF-emphasis plus fast-release single-band compression and pre-compression HF-emphasis plus fast-release 2-band compression were essentially the same. The results did not reveal any remarkable superiority of fast-release 2-band compression over fast-release single-band compression. Note, however, that the speech signal was HF-emphasized before compression. This finding can be easily explained. Pre-compression HF-emphasis has the effect of the equalizing the low-and high-frequency band intensity variations that occur both sequentially as well as at any given moment in time. When fast-release single-band compression is applied after such intensity variations have been controlled, it is essentially acting as a fast-release 2-band compression. Thus, the nature of single-band and 2-band compression used in this study are grossly equivalent because of the pre-compression HF-emphasis. Therefore, fast-release 2-band compression did not prove to be superior to fast-release single-band compression in this study. However, since the extent of HF-emphasis needed to equalize the low- and high-frequency bands is talker-dependent, it can still be argued that fast-release 2-band compression alone should be more effective in reducing the dynamic range of speech than fast-release single-band compression alone.

Comparison of the PI functions of the common conditions between the 3 experiments revealed that they are fundamentally similar. This finding indicates that the 3 groups of subjects are equivalent and that the performance effects demonstrated in this study can be reproduced with a high degree of accuracy.

This finding supports the utility of PI function as a measure of the perceptual dynamic range of speech.

Comparison of fast-release single-band compression with the no-processing revealed that the fast-release single-band compression had an effective compression ratio of 1.6:1 to 1.8:1 over most of the PI function. This finding is superior to HF-emphasis-alone which showed itself to be effective only between the 10% and 60% correct recognition points. However, this effective compression ratio is much lower than that intended. Comparison of fast-release 2-band compression with no-processing revealed a slightly superior effective compression ratio than that provided by fast-release single-band compression.

The fact that the fast-release wide dynamic range compression was found to have some benefit in reducing the overall dynamic range of speech suggests that the issue of amplitude compression as a means of improving speech recognition for severe and profound hearing loss individuals has potential, although lower than that predicted, and that this issue maybe worthy of a re-evaluation.

5.7. SUMMARY

The unprocessed CVC words have an overall perceptual dynamic intensity range of about 30 dB. A simple high-frequency emphasis of 8 dB, intended to equalize the low- and high-frequency bands, does not reduce the overall dynamic range of the speech signal, but enhances perception at low to moderate sensation levels. Between the 10% and 60% correct recognition points, the effective compression ratio of the HF-emphasis-alone was about 1.5:1. When fast-release single-band compression with a compression ratio of 3:1 was applied to the HF-emphasized speech signal, it decreased the overall dynamic range of speech and increased the compression ratio. Over the HF-emphasis-alone condition, the effective compression ratio of the fast-release single-band compression was about 1.5:1 over most of the PI function. This compression ratio was lower than that intended. When compared with no-processing, the effective compression ratio was higher especially at low sensation levels. Between the 10% and 60% correct recognition points, the effective compression ratio was between 1.6:1 and 1.7:1.

Fast-release 2-band compression, with a compression ratio of 3:1 in the low- and high-frequency bands, was not found to be superior over the fast-release single-band compression in reducing the dynamic range of already HF-emphasized speech signal. It is not clear if, in the absence of pre-compression HF-emphasis, the fast-release 2-band compression may be superior to fast-release single-band compression.

This study confirms that fast-release wide dynamic range compression, although not as efficient as intended, does has the effect of reducing the overall dynamic range of speech and enhancing perception at low to moderate sensation levels. This finding calls for renewed efforts to examine the benefit of fast-release wide dynamic range compression as a means of enhancing speech understanding abilities of individuals with severe to profound hearing loss.

CHAPTER VI: GENERAL DISCUSSION AND SUMMARY

6.1. INTERPRETATION OF FINDINGS:

When measured between the 10th and 90th percentiles of the cumulative distribution of RMS amplitude (measured over 125 ms), the unprocessed speech had an intensity range of approximately 30 dB (Figure 3.3). Similarly, the PI function for subjects with normal hearing rose from "toe" to within a few % points of asymptote over a range of approximately 30 dB (Figure 5.4). These findings are in good agreement with each other and with those reported in the literature (e.g., Fletcher, 1953; Boothroyd, 1993; Boothroyd et al., 1994). Taken at face value, the data indicates that an RMS amplitude distribution, using an integration time similar to that of the human ear, provides a reasonable estimate of range over which useful information is distributed in the acoustic speech signal.

A similar conclusion can be drawn from the acoustic and perceptual data for HF-emphasized speech (Figs 3.3 and 5.4). From the perceptual data, however, it is clear that the distribution of useful information is biased towards the higher amplitude levels. This inference is based on the fact that the PI function for HF-emphasized speech is steeper than that for unprocessed speech over the range from 0 to 60% (Figure 5.4). It is also in keeping with expectations. Without HF-emphasis, subjects will hear only low-frequency cues when listening at very low sensation levels. Amplification of the high frequencies will enable them to hear both high- and low-frequency cues as soon as the speech becomes audible. It should not be assumed, therefore, that the shape of the cumulative

RMS amplitude distribution can be used to draw conclusions about the actual distribution of useful speech information across the intensity range.

Because of its effect on slope of the PI function, HF-emphasis provided what was, effectively, compression, but only at low sensation levels. This ratio approached 2:1 between phoneme recognition scores of 10% and 40% (Figure 5.5). Similar results have been reported previously by Boothroyd (1993). In the present study, however, the benefit was lost at higher sensation levels, because the slope of the PI function fell. When measured across the complete PI function the effective compression ratio was zero. This last effect was not observed in the data reported by Boothroyd (1993). It should be noted, however, that Boothroyd (1993) used HF-emphasis at a rate of 4 dB per octave between 500 and 4 KHz whereas the present study involved amplification of all frequencies above 1100 Hz by 8 dB. The topic of spectral modification and its effect on the shape of the PI function clearly warrants further research.

The addition of fast-release single-band compression to HF-emphasis reduced the range of the PI function by about 1.5 to 1 when measured between phoneme recognition scores of 10% and 90% (Figure 5.7). This finding is encouraging in that it provides perceptual validation of the compression algorithm. The magnitude of the effect, however, is only about half of the predicted ratio of 3 to 1. The acoustic data provide a possible explanation of the discrepancy. When measured over an integration time of 125 ms, the RMS amplitude of the high- and low-frequency bands of the compressed signal differed by up to ± 5.5 dB from that predicted by the compression algorithm

(Figure 4.5). Such deviations could easily account for the differences between measured and predicted performance benefits in this and previous studies.

So far, the results of the perceptual study support the qualitative predictions - namely that HF-emphasis-alone will reduce perceptual dynamic range and that fast-release single-band compression will reduce it further. Contrary to prediction, however, the switch from single-band to 2-band fast-release compression produced little or no further reduction of perceptual dynamic range (Figure 5.8). At first sight, this finding seems to be at odds with the acoustic data. The RMS amplitude of the high- and low-frequency bands of the 2-band compressed signal differed by up to only ± 2.5 dB from that predicted by the compression algorithm when measured using 125 ms integration time (Figure 4.5). This apparent conflict may indicate, however, that the dynamic input/output functions used for acoustic validation do not adequately reveal the effectiveness of the compression algorithm. While 2-band compression may reduce the peak-to-trough range within each band, there may be important intensity variations across frequency within each band that are not effectively compressed. 2-band compression is, after all, merely single-band compression within each band. Just as it is argued that the whole spectrum needs to be divided into 2-bands in order to reduce the amplitude difference between two simultaneous cues in the high and low frequencies, so it could be argued that each band needs to be further divided. Such considerations lead towards the conclusion that, to be fully effective, multi-band compression must use a large number of independent bands - perhaps having bandwidths similar to those of the critical bands of the auditory system. As the number of compression bands is increased, however, the effect of compression could be to reduce the very spectral peaks that serve to define phonemic differences. This problem might be especially severe for individuals with sensorineural hearing loss who already exhibit broadened auditory filters (Levitt and Neuman, 1991; Moore, 1991). The acoustic analysis of the present study provides a potential explanation for the failure of fast-release 2-band compression to reduce the perceptual dynamic range of speech below that already accomplished by HF-emphasis and fast-release single-band compression. It is not clear, however, that further increases in the number of bands would improve performance.

Despite the negative findings of the present study regarding the perceptual benefits of fast-release 2-band compression, it would be wrong to conclude that such processing has nothing to offer. It is important to note that, in this study, high-frequency emphasis was applied before the compression. Moreover, the amount of emphasis, was determined from the actual materials to be used in the evaluation of phoneme recognition performance. In the real world, the appropriate amount of high-frequency-emphasis will vary with talker (see, for example, Boothroyd et al., 1994). It would also be expected to vary with talker distance, effort, and orientation, and with the content of the speech (for example, CVC words, versus polysyllabic words, versus sentences). A 2-band compression system, by adjusting average levels in the two bands, would, essentially introduce the appropriate amount of high-frequency emphasis across a variety of conditions. Extension of the present study to include 2-band compression without high-frequency pre-emphasis would provide a test of this prediction.

6.2. THREATS TO INTERNAL VALIDITY:

One of the problems with the use of a word repetition task as a measure of phoneme recognition is the likelihood of errors of interpretation of subjects' utterances. This problem can become especially severe when the subject's percept is of a nonsense syllable, in which case a written response may be difficult to interpret. In the present study, the fact that all subjects had normal hearing reduced the chances of such errors. To reduce them further, the experimenter recorded responses for off-line listening and compared his interpretations with the written responses. To the extent that errors of judgment may still have occurred, the fact that the same judge was used throughout makes it unlikely that the basic findings of the study were affected.

It will be observed from Figures 5.4 and 5.6 that the “toes” of the PI functions for two processing conditions did not always correspond. The discrepancies may be attributed to the difficulty of matching speech levels in a meaningful way when essentially different forms of processing are used. A consequence of this problem is that apparent improvements of performance with processing (as in Figure 5.4) could well reflect a simple shift in average RMS level. A similar argument could be applied to the apparent lack of benefit from fast-release single-band compression at low sensation levels in Figure 5.6. It was to eliminate this threat to internal validity that processing conditions were compared in terms of the dB difference between two percentage scores on the

P1 function rather than in terms of percent correct at two nominally equal presentation levels. In Figure 5.12, however, presentation levels were referenced to the "toe" of the PI function as defined by a phoneme recognition score of 10%. Assuming that the 10% point is an appropriate choice, the curves in Figure 5.12 can be used to compare phoneme recognition performance across conditions as a function of presentation level.

The function used for curve fitting in Figures 5.4, 5.6, and 5.8 contains some arbitrary assumptions about the form of the underlying PI functions. Nevertheless, the resulting curves appear to fit the data very well. Most of the conclusions reached in the present study were based on the use of these curves for interpolation between empirical data points. There is little danger of error in such an approach. A greater threat to validity, however, comes with extrapolation beyond the empirical data points. If, for example, decisions were to be made about the absolute "toe" of the various PI functions, potential errors of extrapolation would become serious. It was for this reason that a phoneme score of 10% was used as the reference criterion in Figure 5.12. Even here, however, the absence of data down to scores of 10% for the HF-emphasis-alone condition raises some concern. Future studies should ensure that performance is measured over as wide a range of presentation levels as possible, in order to avoid the risk of errors of extrapolation.

6.3. THREATS TO EXTERNAL VALIDITY:

The consistency of results within groups and the similarity of group means for identical processing conditions across experiments, lend confidence to the conclusion that the data reported here adequately represent the population means for young normally hearing adults.

It should be noted, however, that these findings were obtained with a particular recording of a particular speaker, producing a rather restricted corpus of spoken utterances. Any generalization to other talkers or other speech materials must, therefore, be done with caution.

Similarly, the extent to which one can generalize these findings to hearing-impaired subjects needs to be explored. It should be noted, in fact, that these experiments were not intended to answer questions about the benefits of compression for hearing-impaired subjects. Normally hearing subjects were used only for perceptual validation of the processing schemes. The results of the present study reveal only the maximum benefit that could be realized by hearing-impaired individuals.

Another issue affecting generalizability is the nature of the compression algorithm used in the present study. A "look-forward" algorithm in which current gain is determined by current input is difficult, if not impossible, to incorporate into an analog hearing aid. Even in digital hearing aids, manufacturers tend to rely on more traditional compression algorithms in which current gain is determined by past input. The present data may not, therefore, generalize to the kinds of hearing aids that are currently available to hearing impaired persons.

Indeed, it is not clear whether a look-forward algorithm is even desirable in a personal hearing aid. It's real-time implementation requires that the output signal be delayed in relation to the input by an amount equal to half of the window used to estimate RMS amplitude. The time delay might be detrimental to hearing-impaired individuals because many of them rely on the simultaneous use of acoustical and visual cues for speech understanding. Note, also, the potential effects of delayed feedback on speech production[*].

Despite these concerns, the results of this study strongly support the idea that, with suitable processing, potential speech perception performance can be enhanced - at low sensation levels, by processing of the type used here. From the curves of Figure 5.12, it may be concluded that, for an individual listening at a sensation level of 8 dB, HF emphasis will raise the maximum possible phoneme recognition score from 50% to 62%; that the addition of fast-release single-band compression will raise it to 77%; and that a switch to fast-release 2-band compression might raise it to 80%. These data do not appear to support a role for fast-release 2-band compression, but, as indicated earlier, fast-release 2-band compression by itself should also provide much of the benefit found here with HF-emphasis followed by fast-release single-band compression.

---

[*] In the algorithm used here, the delay would only be 2.5 ms. Such a delay is probably insignificant but there is no guarantee that a 5 ms integration time is optimal for this application.

6.4. FURTHER RESEARCH:

It was argued earlier that fast-release 2-band compression alone would have produced the results obtained with pre-compression HF-emphasis plus fast-release 2-band compression. Thus, there is an obvious need to explore the effects of fast-release 2-band compression alone - without HF pre-emphasis. Moreover, such studies should include a variety of talkers, recorded under a variety of conditions, especially with changes of average level (to increase long-term dynamic range). The potential negative effects of noise ought also be addressed. The expectation is that fast-release 2-band compression will be able to accommodate talker and other differences better than single-band compression, even with a fixed HF emphasis.

In the present study, a look-forward algorithm was used in which compression was applied based on current input levels. Most compression hearing aids incorporate a look-backward algorithm, in which case compression is applied based on previous input levels. In such algorithms, the effective compression ratio is affected by additional factors such as the attack/release times. As the release time is increased we expect to obtain a reduction of the effective compression ratio. Therefore, the effective compression ratio may be even less with compression algorithms that incorporate look-backward compression than look-forward compression. Thus, perceptual validation procedures are clearly justified for look-backward compression algorithms.

Subjective impressions of a hearing aid can influence its acceptance/rejection by hearing-impaired individuals. Fast-release amplitude compression has the potential to enhance background noise during silent intervals in speech. Further, fast-release amplitude compression introduces distortions of the speech envelope. Such factors may cause a hearing-impaired individual to reject a compression hearing aid even if it enhances speech intelligibility. Therefore, subjective impressions need to be carefully considered in evaluating its benefit for hearing-impaired individuals.

In the present study, only normal-hearing young adults were used. The extent to which the results of this study apply to individuals with sensorineural hearing loss needs to be explored by extending these studies to include actual hearing-impaired subjects.

6.5. CLINICAL IMPLICATIONS:

The present study demonstrated the usefulness of dynamic input/output functions for acoustical evaluation of compression hearing aids. Such methods could be incorporated into hearing aid test protocols to enhance the face validity of acoustical measures of hearing aid performance. Similarly, PI functions can be measured, involving testing with normal-hearing individuals, for perceptual validation of compression schemes to essentially derive the upper limits of possible benefits that could be realized when such schemes are used by individuals with sensorineural hearing loss.

The results of this study revealed that the benefits of wide dynamic range fast-release compression is noted only when listening to speech at low to moderate sensation levels and that the advantage of fast-release compression is nullified when a more or less fully audible speech signal is compressed. This finding would imply that perceptual evaluation of compression schemes in hearing-impaired individuals should involve testing at various sensation levels. Also, the above finding would imply that wide dynamic range fast-release compression would be especially beneficial to individuals who are restricted to listen to speech only at low sensation levels. Indeed such listening restrictions are often encountered by individuals with severe and profound hearing loss due to severely restricted dynamic range of hearing, problems of feedback with high gain aids, etc.

6.6. CONCLUSIONS:

1. The results of the acoustical study support the conclusion that the intensity range of speech can be reduced by fast-release single-band or 2-band compression.
2. The magnitude of the reduction, however, does not match that predicted from the steady state characteristics of the compression algorithm, unless the measuring algorithm uses exactly the same parameters as the compression algorithm.
3. The results of the perceptual study confirm that the intensity range of the useful information in speech is reduced by high-frequency emphasis and further reduced by the addition of fast-release single-band compression.
4. The magnitude of the effect of fast-release single-band compression is less than that predicted. It is, however, in keeping with the results of the acoustical analysis.

5. A change from fast-release single-band to fast-release 2-band compression had little or no effect on the intensity range of the useful information in the speech signal. This finding can be explained if it is assumed that the shortcomings of single-band compression also apply within each band of 2-band compression.
6. Although the potential perceptual benefits of compression fell short of prediction, they were, nevertheless, substantial when expressed in terms of phoneme recognition at low sensation levels.
7. Moreover, there are reasons to predict that fast-release 2-band compression by itself would combine the benefits of high-frequency emphasis and fast-release single-band compression, and that it would do so over a wide range of talker and speech characteristics.
8. The results of this study support the conclusion that PI functions, using normally hearing subjects, can offer effective perceptual validation of amplitude compression schemes before instituting clinical trials on hearing-impaired subjects.
9. The results of this study also support the conclusion that appropriate fast-release compression schemes can increase the potential speech perception performance of hearing-impaired individuals, but only if they must listen all or part of the time at low sensation levels.

APPENDIX A

AB ISO-PHONEMIC WORD LISTS

LIST 1
SHIP
RUG
FAN
CHEEK
HAZE
DICE
BOTH
WELL
JOT
MOVE

LIST 2
FISH
DUCK
PATH
CHEESE
RACE
HIVE
BONE
WEDGE
LOG
TOMB

LIST 3
THUG
WITCH
TEAK
WRAP
VICE
JAIL
HEN
SHOWS
FOOD
BOMB

LIST 4
FUN
WILL
VAT
SHAPE
WREATH
HIDE
GUESS
COMB
CHOOSE
JOB

LIST 5
FIB
THATCH
SUM
HEEL
WIDE
RAKE
GOES
SHOP
VET
JUNE

LIST 6
FILL
CATCH
THUMB
HEAP
WISE
RAVE
GOT
SHOWN
BED
JUICE

LIST 7
BADGE
HUTCH
KILL
THIGHS
WAVE
REAP
FOAM
GOOSE
NOT
SHED

LIST 8
BATH
HUM
DIG
FIVE
WAYS
REACH
JOKE
NOOSE
POT
SHELL

LIST 9
HUSH
GAS
THIN
FAKE
CHIME
WEAVE
JET
ROB
DOPE
LOSE

LIST 10
JUG
LATCH
WICK
FAITH
SIGN
BEEP
HEM
ROD
VOTE
SHOES

LIST 11
MATH
HIP
GUN
RIDE
SIEGE
VEIL
CHOSE
SHOOT
WEB
COUGH

LIST 12
HAVE
WIG
BUFF
MICE
TEETH
JAYS
POACH
RULE
DEN
SHOCK

LIST 13
KISS
BUZZ
HASH
THIEVE
GATE
WIFE
POLE
WRETCH
DODGE
MOON

LIST 14
WISH
DUTCH
JAM
HEATH
LAZE
BIKE
ROVE
PET
FOG
SOON

LIST 15
HUG
DISH
BAN
RAGE
PIES
WET
COVE
LOOSE
MOTH

LIST 16
WAGE
RAG
BEACH
CHEF
DIME
THICK
LOVE
ZONE
HOP
SUIT

LIST 17
JADE
CASH
THIEF
SET
WINE
GIVE
RUB
HOLE
CHOP
ZOOM

LIST 18
SHAVE
JAZZ
THEME
HEIGHT
WIN
SUCK
ROBE
DOG
POOL

LIST 19
VASE
CAB
TEACH
DEATH
NICE
FIG
RUSH
HOPE
LODGE
WOMB

LIST 20
CAVE
RASH
TEASE
JELL
GUIDE
PIN
FUSS
HOME
WATCH
BOOTH

APPENDIX B

DADISP WORKSHEET FOR SIGNAL PROCESSING

This worksheet is used to create stimuli for no-processing, HF-emphasis-alone, single-band compression, and 2-band compression conditions. Listed below are the formulae (and their explanations) used for each window in this worksheet.

W1:

```
readb('f:\ab\ab01\ab0101.wav',3)
```

The original AB word is read in as a signed integer file.

W2:

```
extract(movavg(20*log10(spectrum(hanning(extract(W1,23,2048)))),21),
11, 1024)| setxoffset(0)
```

This formula is used to measure the spectrum of the background noise which is present in all stimuli as a result of room environment. The spectrum is measured in the following way:

1. 2048 points are extracted from the 23rd point in the pre-onset part of the stimuli (each original binary stimulus file has been provided with a 200 ms silent interval at beginning and end of each stimuli). The first 22 points have been excluded from analysis since they contain the header information.
2. A hanning window is then applied on this extracted portion.
3. The spectrum of this extracted portion is obtained.
4. The spectrum is then converted into SPL units (20*log10)
5. A 21 point moving average of the spectrum is then obtained.
6. Finally, 1024 points of the moving average spectrum is extracted form the initial 11th point.

The results of this analysis revealed that the noise is predominately of LF, below 300 Hz.

W3:

```
highpass(22050.0,300.0,3,50.0,1)
```

A high-pass filter is created to eliminate the LF background noise. This filter has a cut-off frequency of 300 Hz, a 3 dB ripple in the passband, and a attenuation of 50 dB down to 1 Hz.

W4: filter(extract(W1/2,23,length(W1)-22),W3)|setdeltax(1/22050)|setxoffset(0)

This formula first removes the header information of the original signal of window w1, and then passes it through the high pass filter of W3. Also, the x-axis units is converted from sampling point to time in seconds.

W5: extract(movavg(20*log10(spectrum(hanning(extract(W4,1,2048)))),21), 11,1024)+6-W2|setxoffset(0)

This formula is used to obtain the low-frequency attenuation characteristics of the low pass filter used to remove the low-frequency noise in the original recordings.

W6: Gline(46,1,1,1)

This formula is used to set the compression threshold. Only that part of signal which is above this threshold gets compressed. For the present study, the compression threshold is 46 dB. It is chosen because, the peak RMS level of the cumulative amplitude distribution of the HF emphasis condition is 75 dB. Thus, approximately the top 30 dB of the signal is being compressed.

W7: Gline(8,1,1,1)

In this window, the amount of HF emphasis is specified. For the present study, a 8 dB boost is given for speech signal above 1100 Hz. The goal of the HF emphasis is to equate the 90th percentile peak levels (composite levels of the entire 20 lists) of the LF and HF bands of the no-processing condition. As figure 4.3 reveals, at the 90th percentile point of the cumulative amplitude distribution for the no-processing condition, the HF band is about 8 dB lower than the LF band.

W8: Gline(20,1,1,1)

This window specifies the amount of post-compression amplification. For the present study a 20 dB post-compression amplification has been chosen so as to equate the 100th percentile peak level of the cumulative amplitude distributions of the HF emphasis alone, single-band and 2-band compression conditions. Further, the 20 dB uniform post-compression gain for all stimuli would essentially reduce the inter-word peak intensity differences.

W9: lowpass(22050.0,950.0,1.0,90.0,1850.0)

A low-pass filter is generated with a cut-off frequency of 950 Hz, 1 dB ripple in the pass band, and an attenuation of 90 dB down to 1850 Hz.

W10: highpass(22050.0,1250.0,1.0,90.0,500.0)

A high-pass filter is generated with a cut-off frequency of 1250 Hz, 1 dB ripple in the pass band, and an attenuation of 90 dB down to 500 Hz.

W11: filter(W4,W9)

The noise-free signal in “window 4” is filtered with the low-pass filter in “window 9”

W12: filter(W4,W10)

The noise-free signal in “window 4” is filtered with the high-pass filter in “window 10”

W13: .01+extract(interpolate(transpose(sqrt(colmean(ravel(W11^2,110,1,55)))), 55),- 56,length(W4))|setdeltax(1/22050)|setxoffset(0)

This formula generates the RMS amplitude envelope of the low-pass filtered signal using 5ms temporal windows, overlapped by 2.5 ms. This is done in the following steps:

1. each data point of the low-pass filtered signal series is first squared.
2. the squared series is then raveled into multiple 110 point (5 ms) segments, where each segment overlaps the previous segment by 55 points (2.5 ms)
3. the square root of the means of the raveled segments is then computed and placed in a row.
4. The row of RMS values of the raveled segments is transposed.
5. Between the existing transposed RMS value points, 55 linear points are added (interpolated)
6. Finally, the interpolated series is extracted to form the temporal envelope of the low-pass filtered signal.

W14: W11/W13

The low-pass filtered signal is divided by its temporal envelope to obtain a unit amplitude signal.

W15: 20*log10(W13)

The temporal envelope of the low-pass filtered signal in “window 13” is converted from linear units into log (sound pressure level) units.

W16: clip(W15,0,length(W6))

The bottom portion of the log envelope of the low-pass filtered signal in "window 15" is clipped based on the compression threshold specified in "window 4". This portion of the temporal envelope (referred henceforth as pedestal) is not compressed

W17: length(W8)+(clip(W15,length(W6),max(W15))-length(W6))/3+W16| overplot(W15)

The top portion of the log envelope of the low-pass filtered signal is clipped and compressed by a factor of 3 (specified in "window 6"). The compressed top potion is then attached to its pedestal. Finally, a 20 dB boost (specified in "window 8") is given to this compressed envelope such that the peaks of the compressed envelope matches to that of its uncompressed counterpart.

W18: 10^(W17/20)

The compressed log envelope of the low-pass filtered signal is converted back into linear units envelope.

W19: W14*W18

The unit amplitude low-pass filtered signal is multiplied by its compressed envelope to obtain the compressed low-pass filtered signal.

W20:

.01+extract(interpolate(transpose(sqrt(colmean(ravel(W12^2,110,1, 55)))),5 5),-56,length(W4))|setdeltax(1/22050)|setxoffset(0)

This formula generates the RMS amplitude envelope of the high-pass filtered signal using 5ms temporal windows, overlapped by 2.5 ms.

W21: W12/W20

The high-pass filtered signal is divided by its temporal envelope to obtain a unit amplitude signal.

W22: 20*log10(W20)+length(W7)

The temporal envelope of the high-pass filtered signal in "window 20" is converted from linear units into log (sound pressure level) units. Also, a 8 dB boost is given to this high-pass filtered signal (specified in "window 7").

W23: clip(W22,0,length(W6))

The pedestal of the log envelope of the high-pass filtered signal in "window 22" is clipped based on the compression threshold specified in "window 4".

W24: length(W8)+(Clip(W22,length(W6),max(W22))-length(W6))/3+W23| overplot(W22)

The top portion of the log envelope of the high-pass filtered signal is clipped and compressed by a factor of 3 (specified in "window 6"). The compressed top potion is then attached to its pedestal. Finally, a 20 dB boost is given to this compressed envelope such that the peaks of the compressed envelope matches to that of its uncompressed counterpart.

W25: 10^(W24/20)

The compressed log envelope of the high-pass filtered signal is converted back into linear units envelope.

W26: W21*W25

The unit amplitude high-pass filtered signal is multiplied by its compressed envelope to obtain the compressed high-pass filtered signal.

W27: extract(W19+W26,1000,length(W4)-2000)

The low-pass and high-pass filtered and compressed signal are added to obtain the "2-band compression with pre-compression HF emphasis signal. Also, 1000 points (45 ms) are deleted from beginning and end of each stimulus. These deleted segments are only the pre-onset and post-offset silent intervals of the speech signal.

W28: W11+W12*10^(length(W7)/20)

The high-pass filtered signal in "window 12" is given a boost of 8 dB (specified in "window 7"). This boosted high-pass filter signal is added with the low-pass filtered signal to derive the "HF emphasis alone" signal.

W29: .01+extract(interpolate(transpose(sqrt(colmean(ravel(W28^2,110,1,55)))), 55),-56,length(W4))|setdeltax(1/22050)|setxoffset(0)

This formula extracts the RMS envelope of the HF emphasized speech signal in "window 28".

W30: W28/W29

The HF emphasized speech signal is divided by its temporal envelope to derive the unit amplitude signal.

W31: 20*log10(W29)

The RMS envelope of the HF emphasized signal is converted from linear units into log units.

W32: clip(W31,0,length(W6))

The pedestal of the log envelope of the HF emphasized signal in "window 31" is clipped based on the compression threshold specified in "window 4".

W33: length(W8)+(Clip(W31,length(W6),max(W31))-length(W6))/3+W32| overplot(W31)

The top portion of the log envelope of the HF emphasized signal is clipped and compressed by a factor of 3 (specified in "window 6"). The compressed top potion is then attached to its pedestal. Finally, a 20 dB boost is given to this compressed envelope.

W34: 10^(W33/20)

The compressed envelope of the HF emphasized signal is converted from log units into linear units.

W35: extract(W34*W30,1000,length(W4)2000)|overplot(gline(length(w4), 1/22050,0,2^15))|overplot(gline(length(w4),1/22050,0,-^15)))| overplot(gline(length(w4),1/22050,0,-2^15))

The compressed envelope of the HF emphasized signal is multiplied by its unit amplitude waveform to obtain the "1-band compression with pre-compression HF emphasis" signal.

W36: w27|overplot(gline(length(w4),1/22050,0,2^15))|overplot(gline(length(w4), 1/22050,0,-2^15))|overplot(gline(length(w4),1/22050,0,-2^15))

The "2-band compression with pre-compression HF emphasis" signal from window 27 gets copied to this window.

W37: `extract(w28,1000,length(w4)2000)|overplot(gline(length(w4),1/22050,0,2^15))|overplot(gline(length(w4),1/22050,0,2^15))|overplot(gline(length(w4),1/22050,0,-2^15))`

The "HF emphasis alone" signal from "window 28" gets copied to this window. Also, 1000 points from the pre-signal onset and post-signal off-set portion gets deleted.

W38: `extract(extract(W4,1000,length(W4)-2000),-1000,length(w4))`

The LF noise free signal in "window 4" gets copied to this window. Also, 1000 points from the pre-signal onset and post-signal off-set portion gets deleted.

APPENDIX C

MEASUREMENT OF CUMULATIVE AMPLITUDE DISTRIBUTION

Described below is the measurement procedures for obtaining mean RMS cumulative amplitude distributions using DaDisp. Measurement was completed in 2 stages. In the first stage, the cumulative distribution of 10 word stimuli of a given list was calculated. In the second stage, the cumulative distribution of the 20 lists of the words were calculated.

Stage 1: Worksheet "ABLIST"

The 10 word stimuli of a given word list was imported as signed integer. The silent intervals present at the pre-onset and post-offset of each stimulus was deleted. Then, the 10 stimuli were concatenated and filtered through a low-pass filter and a high-pass filter. The filter characteristics were same as were used for signal processing in this study. The RMS cumulative amplitude distributions of the full-band, low-pass and high-pass filtered stimuli were then calculated using a temporal integration time of 125 ms. This process was repeated for each of the 20 AB word lists.

Stage 2: Worksheet "Cumdist"

To obtain the mean RMS cumulative amplitude distribution of the entire 20 lists, the amplitude distributions of the 20 individual lists were concatenated, raveled, and then averaged.

Worksheet "ABLIST"

W 1 through W10: e.g. readb('e:\abnf\ab0101nf.wav',3)

The 10 CVC words of a given list of no processing condition were imported as a signed integer/binary file.

W 11 through W20: e.g. w11: extract(W1,4433,length(W1)-8842)

Each stimulus had 200 ms pre-onset and post-offset silent intervals - these silent intervals were deleted.

W 21: Concat(W11..w20)

The 10 extracted stimuli were then concatenated.

W 22: ampdist(concat(20*log10(transpose(sqrt(colmean(Ravel(W21^2,
2756,1,1378))))),gseries(0)),1)

This window calculated the full-band amplitude distribution of the concatenated string of stimuli. First, each data point in the concatenated string was rectified by squaring each of its points and then raveled into segments 125 ms long. Each segment overlapped the previous segment by 62.5 ms (i.e., half the duration of the window). The square root of the means of columns of raveled segments were calculated and transposed into a single row. These transposed data were then converted into decibel units (20*log10). The result represented RMS level as a function of time. Finally, the amplitude distribution of the RMS function was calculated.

W 23: Integ(W22)/max(integ(W22))*100

The full-band amplitude distribution was integrated, divided by the maximum integrated value and multiplied by 100. Finally, the amplitude distribution data were plotted as percent cumulative energy of the list as a function of intensity.

W 24: FILTER.1.LOWPASS

This window contains the low-pass filter used in the signal processing schemes of the present study.

W 25: Filter(W21,W24)

The concatenated string of stimuli in window 21 was filtered through the low-pass filter.

W 26: `ampdist(concat(20*log10(transpose(sqrt(colmean(Ravel(W25^2, 2756,1,1378))))),gseries(0)),1)`

The amplitude distribution of the low-frequency band of the concatenated string is calculated.

W 27: `integ(W26)/max(integ(W26))*100`

This window shows the mean cumulative amplitude distribution of the low-frequency band of 1 AB word list.

W 28: `FILTER.1.HIGHPASS`

This window contains the high-pass filter used in the signal processing schemes of the present study.

W 29: `Filter(W21,W28)`

The concatenated string of stimuli in window 21 was filtered through the high-pass filter.

W 30: `ampdist(concat(20*log10(transpose(sqrt(colmean(Ravel(W29^2, 2756,1,1378))))),gseries(0)),1)`

The amplitude distribution of the high-frequency band of the concatenated string is calculated.

W 31: `Integ(W30)/max(integ(W30))*100`

This window shows the mean cumulative amplitude distribution of the high-frequency band of 1 AB word list.

Finally, the datum in windows 23, 27, and 31 were saved as dataseries.

Worksheet "Cumdist"

W 1 through W20: e.g. W1: CUMDISTUP.1.01

The cumulative amplitude distribution of the 10 words of a list of a given condition and frequency band is opened.

W 21 through W40: e.g. W21: -extract(-w1+100,1,80)+100

The dB range of the x-axis was set from 0 to 80 dB. A horizontal straight line was drawn from the dB point representing the 100th percentile point of the cumulative amplitude distribution to the upper extent of the dB range of the x-axis.

W 41: Ravel(concat(w21..w40),80,1)|overplot(w42)|sety(0,104)|setytic(10)| setxtic(1)

The 20 lists were concatenated and raveled into 80 point segments.

W 42: Sums(w21..w40)/20

The mean distribution of the entire 20 lists was calculated.

W 43: W41|overplot(w42)|sety(80,90)|setx(68,73)|setxtic(1)

This window plots the average as well as the individual cumulative amplitude distributions of the 20 AB word lists.

APPENDIX D

MEASUREMENT OF DYNAMIC INPUT/OUTPUT FUNCTIONS

Within a DaDisp worksheet, all 10 words of AB word list 1 of a given condition were imported as signed integer/binary files. The pre-onset and post-offset silent intervals present in each stimuli were deleted. The extracted stimuli were then concatenated and filtered through the low-pass and the high-pass filters used in the signal processing. The amplitude distributions of the unfiltered, low-pass filtered, and high-pass filtered concatenated strings were then computed.

The concatenated string was raveled into segments of duration equivalent to the temporal integration time of the measure. For the 5 ms integration time measure, the raveled segments were not overlapped. For the 125 ms integration time measure, however, the raveled segments were overlapped by about 93.75 ms (3/4th of the integration time). Second, the standard deviation of the columns of raveled segments were computed. Finally, the standard deviations were transposed and converted into decibels (20*log10) to obtain the amplitude distribution.

This process was carried out separately for the HF emphasis, single-band compression and 2-band compression conditions. To obtain the input/output functions for a specific frequency band as well as a specific temporal integration window for a given compression condition, an XY plot was drawn showing amplitude distribution of the HF emphasis alone condition on the x-axis and the corresponding amplitude distribution of the compression condition on the Y-axis.

Worksheet "INOUTPUT"

W 1 through W10: e.g. ReadB('f:\ABC1\AB0101C1.WAV',3)

The 10 CVC words of AB word list 1 of HF emphasis alone or single-band or 2-band compression conditions were imported as signed integer/binary files

W 11 through W20: e.g. Extract(W1,4432,length(W1)-8864)

Each stimuli had a 200 ms pre-onset and post-offset silent intervals - these silent intervals were deleted.

W 21: Concat(W11..w20)

The 10 extracted stimuli were then concatenated. Concatenated strings of the HF emphasis alone, single-band and 2-band compression conditions were saved as data series.

W 22: FILTER.1.HIGHPASS or FILTER.1.LOWPASS

The digital high-pass or the low-pass filter used in the signal processing was imported into this window.

W 23: Filter(W21,W22)

The concatenated string of the single-band or the 2-band compression condition words was filtered through the high-pass or the low-pass filter.

W 24: ABHE.1.LIST1

The already concatenated string of HF emphasis alone condition (AB word list 1) was imported into this window.

W 25: Filter(W24,W22)

The concatenated string of the HF emphasis alone condition was filtered through the high-pass or the low-pass filter.

W 26: 20*log10(transpose(colstdev(ravel(W21,110))))

The amplitude distribution of the whole band of single-band or 2-band compression conditions were obtained. In this example, the amplitude distribution for single-band compression was obtained using a temporal integration time of 5 ms. To obtain the amplitude distribution for 125 ms integration time, the concatenated strings were raveled into segments 2756 points long.

W 27: 20*log10(transpose(colstdev(ravel(W24,110))))

The amplitude distribution of the whole band of the HF emphasis alone condition was obtained.

W 28: Xyplot(W27,W26)

The input/output function for the whole band for single-band or 2-band compression was obtained for 5 ms integration time. In this example, the input/output function was obtained for single-band compression.

W 29: 20*log10(transpose(colstdev(ravel(W23,110))))

The amplitude distribution of the low-pass or high-pass filtered string of single-band or 2-band compression conditions were obtained. In this example, the amplitude distribution for single-band compression was obtained using a temporal integration time of 5 ms.

W 30: 20*log10(transpose(colstdev(ravel(W25,110))))

The amplitude distribution of the low-pass or high-pass filtered string of HF emphasis condition was obtained. In this example, the amplitude distribution was obtained using a temporal integration time of 5 ms.

W 31: Xyplot(W29,W30)

The input/output function for the LF or HF band for single-band or 2-band compression was obtained for 5 ms integration time. In this example, the input/output function was obtained for single-band compression.

APPENDIX E

INFORMED SUBJECT CONSENT FORM

Title of the project: The effects of high-frequency emphasis and digital syllabic compression on the short-term dynamic intensity range of speech

Principal Investigator: Balaji Oruganti

The general goal of this study is to evaluate the effect of signal processing on the short-term intensity range of speech. As part of this study, you will be listening to naturally spoken and digitally processed words, in a background of noise. Your task is to write and repeat what you heard. Some of the words will be barely audible while others will be clearly audible. Under no circumstances will you be required to listen to sounds that are known to be potentially dangerous. Your verbal responses will be audio recorded for off-line scoring and analysis. If you wish, you can review the completed audio tape. You will also receive a basic hearing test in a separate session. The approximate duration of this study is about 2 hours. While the results of this study will be reported at professional meetings and may be published, subject anonymity will be protected.

Please note that the Graduate school has a committee which safeguards your interests. This committee requires us to point out that the research done here is of an experimental nature and is done in a university and not a clinical setting. If you have any questions regarding your participation as an experimental subject, please feel free to ask me, or contact either my advisor, Prof. A. Boothroyd [Tel: (212) 642-2352] or the office of sponsored research at the graduate school [Tel: (212) 790-4366].

Sincerely,

Balaji Oruganti
Doctoral student
Ph.D. Program in Speech and hearing Sciences
Graduate school, City University of New York
33 W. 42 Street, New York, NY 10036.

I agree to participate in this experimental investigation described above. I understand that my participation is voluntary and that I may withdraw at any time without penalty.
Name: ____________________________________

Signature: ________________________________

Date: _____________________________________

APPENDIX F

Table F.1: Table F.1a shows the output accuracy and table F.1b shows the output linearity of the GSI-10 clinical audiometer used in the present study. All measurements were made with TDH-50p headphone attached to a 6 cc coupler, and B&K sound level meter (model 2203) attached to a 1 inch microphone (serial number 4144).

a) Output accuracy

Frequency in Hz	ANSI TDH 49 and 50	Variations from normal @ 70 dB HL	
		Left	Right
125	117.6	117.5	117.0
250	96.7	97.0	97.0
500	83.5	83.5	84.0
1000	77.4	77.5	77.5
2000	81.1	81.0	81.0
4000	80.7	81.0	80.5
8000	83.0	83.0	83.0

b) Attenuator linearity at 1000 Hz

Dial Setting	Channel 1	Channel 2
110.0	117.5	117.5
105.0	112.5	112.5
100.0	107.5	107.5
95.0	102.5	102.5
90.0	98.0	98.0
85.0	92.5	92.5
80.0	87.5	87.5
75.0	82.5	82.5
70.0	77.5	77.5
65.0	72.5	72.5
60.0	67.5	67.0
55.0	62.5	62.0
50.0	57.5	57.0
45.0	52.5	52.0
40.0	47.0	47.0
35.0	42.5	42.0
30.0	37.5	37.0
25.0	32.5	32.0
20.0	27.5	27.0
15.0	23.0	22.0
10.0	19.0	19.0
5.0	15.5	16.0
0.0	14.0	14.0

APPENDIX G

Table G.1: Individual subject data of pure tone thresholds in a) quiet, and b) pink noise for experiment 1

a) Thresholds in Quiet

Subj	Sex	Age in yrs	Test ear	Frequency in Hz							
				250	500	1000	2000	3000	4000	6000	8000
				SPL	SPL	SPL	SPL	SPL	SPL	SPL	SPL
S1	F	58	R	27.0	14.0	7.5	26.0	14.7	25.5	33.4	33.0
S2	F	35	L	32.0	23.5	17.5	16.0	9.7	11.0	23.4	23.0
S3	M	33	R	37.0	19.0	7.5	6.0	4.7	20.5	18.4	18.0
S4	M	28	L	37.0	18.5	12.5	11.0	4.7	6.0	13.4	28.0
S5	F	31	R	27.0	19.0	7.5	21.0	19.7	15.5	13.4	23.0
S6	M	24	L	32.0	23.5	17.5	21.0	9.7	11.0	28.4	18.0
S7	F	27	R	32.0	14.0	12.5	21.0	19.7	5.5	23.4	18.0
S8	F	30	L	37.0	18.5	17.5	16.0	19.7	11.0	23.4	18.0
Mean		33.3		32.6	18.8	12.5	17.3	12.8	13.3	22.2	22.4
SD		10.6		4.2	3.6	4.6	6.4	6.5	6.9	6.9	5.6

b) Thresholds in pink noise

Subj	Sex	Age in yrs	Test ear	Frequency in Hz							
				250	500	1000	2000	3000	4000	6000	8000
				SPL	SPL	SPL	SPL	SPL	SPL	SPL	SPL
S1	F	58	R	47.0	44.0	42.5	41.0	39.7	40.5	43.4	38.0
S2	F	35	L	47.0	38.5	37.5	36.0	39.7	36.0	38.4	33.0
S3	M	33	R	52.0	44.0	42.5	41.0	39.7	40.5	38.4	28.0
S4	M	28	L	47.0	38.5	42.5	41.0	39.7	36.0	38.4	38.0
S5	F	31	R	47.0	39.0	37.5	41.0	34.7	40.5	38.4	33.0
S6	M	24	L	52.0	48.5	42.5	41.0	39.7	36.0	38.4	28.0
S7	F	27	R	47.0	44.0	37.5	46.0	34.7	35.5	43.4	33.0
S8	F	30	L	52.0	38.5	42.5	41.0	34.7	41.0	38.4	28.0
Mean		33.3		48.9	41.9	40.6	41.0	37.8	38.3	39.7	32.4
SD		10.6		2.6	3.8	2.6	2.7	2.6	2.5	2.3	4.2

Table G.2: Individual subject data of pure tone thresholds in a) quiet, and b) pink noise for experiment 2

a) Thresholds in Quiet

Subj	Sex	Age in years	Test ear	Frequency in Hz							
				250	500	1000	2000	3000	4000	6000	8000
				SPL	SPL	SPL	SPL	SPL	SPL	SPL	SPL
S9	M	20	R	32.0	19.0	17.5	11.0	9.7	15.5	13.4	18.0
S10	F	40	L	37.0	13.5	2.5	21.0	19.7	16.0	18.4	23.0
S11	M	35	R	42.0	24.0	22.5	21.0	19.7	20.5	18.4	18.0
S12	F	22	L	37.0	18.5	7.5	6.0	14.7	21.0	23.4	18.0
S13	F	33	R	32.0	14.0	12.5	16.0	-0.3	10.5	23.4	13.0
S14	M	25	L	27.0	8.5	7.5	11.0	9.7	6.0	23.4	8.0
S15	F	29	R	32.0	19.0	7.5	16.0	9.7	20.5	13.4	13.0
S16	M	28	L	42.0	23.5	12.5	16.0	14.7	21.0	13.4	3.0
Mean		29.0		35.1	17.5	11.3	14.8	12.2	16.4	18.4	14.3
SD		6.8		5.3	5.3	6.4	5.2	6.5	5.6	4.6	6.4

b) Thresholds in pink noise

Subj	Sex	Age in years	Test ear	Frequency in Hz							
				250	500	1000	2000	3000	4000	6000	8000
				SPL	SPL	SPL	SPL	SPL	SPL	SPL	SPL
S9	M	20	R	52.0	49.0	42.5	41.0	39.7	40.5	38.4	33.0
S10	F	40	L	52.0	43.5	37.5	41.0	39.7	36.0	38.4	33.0
S11	M	35	R	52.0	49.0	42.5	41.0	39.7	35.5	33.4	28.0
S12	F	22	L	52.0	43.5	42.5	41.0	39.7	36.0	38.4	33.0
S13	F	33	R	47.0	44.0	37.5	41.0	39.7	35.5	33.4	28.0
S14	M	25	L	47.0	43.5	42.5	36.0	34.7	36.0	38.4	28.0
S15	F	29	R	47.0	39.0	42.5	41.0	34.7	35.5	38.4	28.0
S16	M	28	L	52.0	43.5	42.5	36.0	34.7	36.0	33.4	28.0
Mean		29.0		50.1	44.4	41.3	39.8	37.8	36.4	36.5	29.9
SD		6.8		2.6	3.3	2.3	2.3	2.6	1.7	2.6	2.6

Table G.3: Individual subject data of pure tone thresholds in a) quiet, and b) pink noise for experiment 2

a) Thresholds in Quiet

				Frequency in Hz							
		Age in	Test	250	500	1000	2000	3000	4000	6000	8000
Subj	Sex	years	ear	SPL	SPL	SPL	SPL	SPL	SPL	SPL	SPL
S17	F	44	R	37.0	24.0	17.5	16.0	4.7	10.5	23.4	18.0
S18	M	31	L	37.0	13.5	12.5	6.0	19.7	21.0	18.4	18.0
S19	F	31	R	27.0	19.0	12.5	11.0	9.7	15.5	23.4	23.0
S20	M	26	L	27.0	13.5	7.5	11.0	9.7	16.0	18.4	18.0
S21	M	44	R	37.0	24.0	17.5	16.0	14.7	15.5	13.4	18.0
S22	F	24	L	37.0	13.5	2.5	6.0	-0.3	1.0	18.4	23.0
S23	M	27	R	37.0	24.0	12.5	16.0	9.7	15.5	23.4	23.0
S24	F	37	L	32.0	23.5	7.5	6.0	4.7	11.0	18.4	23.0
Mean		33.0		33.9	19.4	11.3	11.0	9.1	13.3	19.7	20.5
SD		7.9		4.6	5.1	5.2	4.6	6.2	5.9	3.5	2.7

b) Thresholds in pink noise

				Frequency in Hz							
		Age in	Test	250	500	1000	2000	3000	4000	6000	8000
Subj	Sex	years	ear	SPL	SPL	SPL	SPL	SPL	SPL	SPL	SPL
S17	M	44	R	52.0	49.0	42.5	41.0	39.7	40.5	38.4	33.0
S18	F	31	L	47.0	48.5	42.5	41.0	39.7	36.0	33.4	33.0
S19	M	31	R	52.0	44.0	42.5	41.0	39.7	40.5	38.4	33.0
S20	F	26	L	47.0	43.5	37.5	41.0	39.7	36.0	33.4	33.0
S21	F	44	R	52.0	49.0	47.5	46.0	44.7	40.5	38.4	33.0
S22	M	24	L	47.0	43.5	37.5	41.0	39.7	41.0	38.4	33.0
S23	F	27	R	47.0	44.0	42.5	41.0	39.7	35.5	38.4	33.0
S24	M	37	L	52.0	48.5	42.5	41.0	39.7	41.0	38.4	28.0
Mean		33.0		49.5	46.3	41.9	41.6	40.3	38.9	37.2	32.4
SD		7.9		2.7	2.7	3.2	1.8	1.8	2.5	2.3	1.8

APPENDIX H

Table H.1: One-third octave band SPL level of the pink noise used in the present study along with variations in output that occurred when pure tone was selected as input for channel 1 of the audiometer and the adjusted attenuator setting of the channel 2 needed for masked threshold determination. These measurements were made at the output of TDH-50p headphone attached to a 6 cc coupler.

Freq. in Hz	output level in dB SPL with speech as input to channel 1	output level in dB SPL with tone as input to channel 1	channel 2 setting (in dB HL) of pink noise used to obtain masked thresholds
250	36.0	39.0	38.0
500	42.0	30.0	52.0
1000	46.0	58.0	58.0
2000	46.0	32.5	53.0
3150	46.0	32.5	54.0
4000	44.5	32.5	52.0
6300	40.0	31.5	48.0
8000	33.0	32.5	40.0

Table I.3: Number of phonemes recoginsed in 1 AB word list (total = 30 phonemes) by each subject at each presentation level and replication forp re-compression HF-emphasis plus single-band compression and pre-compression HF-emphasis plus 2-band compression conditions

		HF-Emphasis plus Single-Band Compression														HF-Emphasis plus 2-Band Compression													
		Presentation Levels														Presentation Levels													
		Replication 1							Replication 2							Replication 1							Replication 2						
Sub	Ear	40a	36a	32a	28a	26a	24a	22a	40b	36b	32b	28b	26b	24b	22b	40a	36a	32a	28a	26a	24a	22a	40b	36b	32b	28b	26b	24b	22b
S1	R	30	28	30	20	18	10	7	30	27	30	21	17	12	7	29	28	26	21	17	15	6	28	30	28	27	20	13	4
S2	L	30	27	28	19	19	9	8	28	30	25	15	15	10	2	30	26	26	25	17	10	3	29	30	26	21	17	13	9
S3	R	30	30	25	21	20	13	1	30	29	27	27	15	14	5	30	29	26	22	15	18	3	30	29	27	26	15	12	6
S4	L	29	29	29	22	13	11	3	29	26	29	20	13	5	7	30	30	23	19	15	16	5	30	29	24	21	16	6	4
S5	R	29	30	28	20	13	15	0	29	30	23	23	15	13	1	30	29	28	20	18	15	8	30	27	29	24	17	11	1
S6	L	29	29	29	22	18	14	8	30	29	28	26	20	16	7	30	29	26	24	18	9	6	30	30	28	24	16	11	6
S7	R	30	28	25	21	13	9	5	29	26	26	24	19	14	4	29	27	29	27	14	12	4	29	29	26	18	23	10	4
S8	L	29	28	25	23	15	11	3	30	29	28	19	12	8	6	26	28	28	24	18	8	5	30	29	26	22	19	13	4

Table I.2: Number of phonemes recoginsed in 1 AB word list (total = 30 phonemes) by each subject at each presentation level and replication for HF-emphasise-alone and pre-compression HF-emphasis plus single-band compression conditions

		HF-Emphasis-Alone														HF-Emphasis plus Single-Band Compression													
		Presentation Levels														Presentation Levels													
		Replication 1							Replication 2							Replication 1							Replication 2						
Sub	Ear	45a	38a	34a	30a	26a	23a	20a	45b	38b	34b	30b	26b	23b	20b	45a	38a	34a	30a	26a	23a	20a	45b	38b	34b	30b	26b	23b	20b
S1	R	30	30	25	25	15	7	5	26	28	26	18	19	9	4	30	29	28	26	23	12	9	30	30	27	27	16	5	5
S2	L	30	26	28	18	14	7	5	28	26	27	22	14	8	5	28	30	25	24	17	8	1	30	28	26	24	18	6	3
S3	R	27	25	23	19	20	9	5	28	29	24	20	10	8	3	29	28	28	25	19	8	3	28	30	26	26	14	7	3
S4	L	30	30	27	22	18	10	4	28	29	22	21	16	12	2	30	27	29	26	22	5	3	30	27	29	19	12	9	2
S5	R	29	30	29	27	11	8	3	30	27	21	18	14	10	1	30	30	26	26	25	8	0	30	28	27	23	18	7	2
S6	L	29	28	27	22	17	13	1	29	25	24	25	17	10	4	30	30	23	23	19	8	1	30	28	27	23	18	8	2
S7	R	30	22	27	22	21	12	4	30	27	26	24	18	8	2	30	29	29	26	22	9	1	30	30	28	26	14	6	0
S8	L	30	26	26	19	15	9	5	30	26	22	23	11	6	2	30	29	25	24	15	7	1	28	28	23	25	10	6	2

Table I.1: Number of phonemes recognised in 1 AB word list (total = 30 phonemes) by each subject at each presentation level and replication for no-processing and HF-emphasis-alone conditions

		No-Processing														HF-Emphasis-Alone													
		Presentation Levels														Presentation Levels													
		Replication 1							Replication 2							Replication 1							Replication 2						
Sub	**Ear**	55a	45a	35a	30a	27a	24a	22a	55b	45b	35b	30b	27b	24b	22b	55a	45a	35a	30a	27a	24a	22a	55b	45b	35b	30b	27b	24b	22b
S1	R	29	26	24	22	12	5	1	30	27	22	22	6	4	3	30	27	24	22	20	10	5	30	28	26	20	14	10	7
S2	L	30	26	19	12	10	6	4	29	29	21	18	10	9	4	30	30	27	21	27	5	9	29	28	28	20	17	16	6
S3	R	30	29	22	17	9	6	7	30	26	24	20	8	11	2	30	30	29	20	18	12	8	30	28	28	24	21	11	7
S4	L	30	29	29	19	8	3	4	30	28	24	11	14	6	3	30	30	23	20	19	12	7	30	26	24	24	12	1	6
S5	R	29	30	19	13	7	10	5	30	29	21	14	15	7	2	30	30	24	15	16	11	11	30	30	26	27	16	8	2
S6	L	30	28	28	11	15	7	2	30	29	26	20	7	5	2	30	29	23	24	17	12	9	30	29	25	21	19	6	8
S7	R	30	27	26	17	12	4	7	28	27	24	12	6	4	7	29	29	27	20	20	15	3	29	28	27	24	18	16	6
S8	L	30	27	21	12	16	4	1	30	29	22	17	9	2	5	30	29	28	17	18	11	6	30	30	24	19	19	12	6

REFERENCES

Abramovitz, R. (1980). Frequency shaping and multiband compression in hearing aids. Journal of Communication Disorders, 13, 483-488.

Allen, J. B., Hall, J. L., and Jeng, P. S. (1990). Loudness growth in 1/2 octave bands (LGOB) - A procedure for the assessment of loudness. Journal of the Acoustical Society of America, 88, 745-753.

American National Standards Institute (1969). Methods for the calculation of the articulation index. S3.5. New York: Acoustical Society of America.

Barfod, J. (1976). Multi channel compression hearing aids. Report # 11. The Acoustical Laboratory, Technical University of Denmark.

Benson, D., Clark, T. M., and Johnson, J. S. (1992). Patient's experiences with multiband full dynamic range compression. Ear and Hearing, 13, 320-330.

Benson, R. W., and Hirsh, I. J. (1953). Some variables in audio spectrometry. Journal of the Acoustical Society of America, 25, 499-505.

Boothroyd, A. (1968). Statistical theory of the speech discrimination score. Journal of the Acoustical Society of America, 43, 362-367.

Boothroyd, A. (1993). Speech perception, sensorineural hearing loss, and hearing aids. In G. Studebaker and I. Hochberg (Eds.), Acoustical factors affecting hearing aid performance (pp. 277-299). Boston: Allyn and Bacon.

Boothroyd, A., Erickson, F. N., and Medwetsky, L. (1994). The hearing aid input: A phonemic approach to assessing the spectral distribution of speech. Ear and Hearing, 15, 1-11.

Boothroyd, A., Springer, N., Smith, L., and Schulman, J. (1988). Amplitude compression and profound hearing loss. Journal of Speech and Hearing Research, 31, 362-376.

Braida, L. D., Durlach, N. I., Lippman, R. P., Hicks, B. L., Rabinowitz, W. M., and Reed, C. M. (1979). Hearing aids: A review of past research of linear amplification, amplitude compression and frequency lowering. ASHA Monograph, 19. Rockville, MD: American Speech-Language-Hearing Association.

Brand, J. F., Ruder, K. F., and Shipp, T. (1969). Vocal loudness and effort in continuous speech. Journal of the Acoustical Society of America, 46, 1543-1548.

Byrne, D. (1977). The speech spectrum - some aspects of its significance for hearing aid selection and evaluation. British Journal of Audiology, 11, 40-46.
Cornelisse, L. E., Gagne, J. P., and Seewald, R. C. (1991). Ear level recordings of the long-term average spectrum of speech. Ear and Hearing, 12, 47-54.

Cox, R. M., Matesich, J. S., and Moore, J. N. (1988). Distribution of short-term RMS levels in conversational speech. Journal of the Acoustical Society of America, 84, 1100-1104.

Cox, R. M., and Moore, J. N. (1988). Composite speech spectrum for hearing aid gain prescriptions. Journal of Speech and Hearing Research, 31, 102-107.

Davis, H., and Silverman, S. R. (1960). Hearing and deafness. 4th edition, New York: Holt, Rinehart, and Winston.

De Gennaro, S., Braida, L. D., and Durlach, N. I. (1986). Multichannel syllabic compression for severely impaired listeners. Journal of Rehabilitation Research and Development, 23, 17-24.

Dillon, H. (1996). Compression? Yes, but for low or high frequencies, for low or high intensities, and with what response times? Ear and Hearing, 17, 287-307.

Duffy, J. R., and Giolas, T. G. (1974). Sentence intelligibility as a function of key word selection. Journal of Speech and Hearing Research, 17, 631-637.

Dunn, H. K., and Farnsworth, D. W. (1939). Exploration of pressure field around the human head during speech. Journal of the Acoustical Society of America, 10, 184-199.

Dunn, H. K., and White, S. D. (1940). Statistical measurements on conversational speech. Journal of the Acoustical Society of America, 11, 278-288.

Elpern, B. (1961). The relative stability of half list and full list discrimination tasks. Laryngoscope, 71, 30-36.

Epstein, A., Giolas, T. C., and Owens, E. (1968). Familiarity and intelligibility of monosyllabic word lists. Journal of Speech and Hearing Research, 11, 435-438.

Fletcher, H. (1953). Speech and Hearing in communication. New York: Van Nostrand.

Formby, C., and Monsen, R. B. (1982). Long-term average speech spectra for normal and hearing-impaired adolescents. Journal of the Acoustical Society of America, 71, 196-202.

Fowler, E. P. (1936). A method for early detection of otosclerosis. Archives of Otolaryngology, 24, 731-741.
Gelfand, S. A. (1998). Optimizing the reliability of speech recognition scores. Journal of Speech, Language, and Hearing Research, 41, 1088-1102.

Gerber, S. E. (1974). Introductory hearing science: physical and psychological concepts. Philadelphia: W. B. Saunders Company.

Giolas, T. G., Cooker, H. S., and Duffy, J. R. (1970). The predictability of words in sentences. Journal of Auditory Research, 10, 328-334.

Hagerman, B. (1976). Reliability in the determination of speech discrimination. Scandinavian Audiology, 5, 219-228.

Hallpike, C. S. (1967). The loudness recruitment phenomenon: A clinical contribution to the neurology of hearing. In Graham, A. B. (Ed.), Sensori-neural hearing processes and disorders (pp. 489-499). Boston: Little and Brown.

Hallpike, C. S., and Hood, J. P. (1959): Observations upon the neurological mechanism of the loudness recruitment phenomenon. Acta Oto-laryngological, 50, 472-486.

Harris, J. D. (1953). A brief critical review of loudness recruitment. Psychological Bulletin, 50,190-203.

Hellman, R. P., and Meiselman, C. H. (1990). Loudness relations for individuals and groups in normal and impaired hearing. Journal of the Acoustical Society of America, 88, 2596-2606.

Hellman, R. P., and Meiselman, C. H. (1993). Rate of loudness growth for puretones in normal and impaired hearing. Journal of the Acoustical Society of America, 93, 966-975.

Jerger, J. (1962). Hearing tests in otological diagnosis. ASHA, 4, 139-143.

Jirsa, R. E., Hodgson, W. R., and Goetzinger, C. P. (1975). Unreliability of half-list discrimination tests. Journal of American Audiology Society, 1, 47-49.

Knight, K. K., and Margolis, R. H. (1984). Magnitude estimation of loudness: Loudness perception in presbycusis listeners. Journal of Speech and Hearing Research, 27, 28-32.

Krieg, K. R. (1980). Third octave band level distributions of amplitude compressed speech. Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

Kuk, F. (1996). Theoretical and practical considerations in compression hearing aids. In M. Valente (Ed), Trends in amplification. New York: Woodland Publishers, Inc.

Launer, S. (1995). Loudness perception by listeners with sensori-neural hearing impairment. Doctoral thesis, Department of physics, University of Oldenburg, Germany.

Laurence, R., Moore, B., and Glasberg, B. (1983). A comparison of behind-the-ear high-fidelity linear hearing aids and 2-channel compression aids in the laboratory and in everyday life. British Journal of Audiology, 17, 31-48.

Levitt, H., and Neuman, A. C. (1991). Evaluation of orthogonal polynomial compression. Journal of the Acoustical Society of America, 90, 241-252.

Lippman, R., Braida, L., and Durlach, N. I. (1981). Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss. Journal of the Acoustical Society of America, 69, 524-534.

Medwetsky, L., and Boothroyd, A. (1991). Effect of microphone placement on the spectral distribution of speech. Paper presented at the annual American Speech-Language Hearing Association conference, Atlanta, Georgia.

Miskolczy-Fodor, F. (1960). Relations between loudness and duration of tone pulses. III. Response in cases of abnormal loudness function. Journal of the Acoustical Society of America, 32, 486-492.

Moore, B. C. J. (1987): Design and evaluation of two-channel compression hearing aids. Journal of Rehabilitative Research and Development, 24, 181-192.

Moore, B. C. J. (1989). An introduction to the psychology of hearing. 3rd edition, London: Academic Press.

Moore, B. C. J., and Glasberg, B. R. (1986). A comparison of two-channel and single-channel compression hearing aids. Audiology, 25, 210-226.

Moore, B. C. J, and Glasberg, B. R. (1988). A comparison of four methods of implementing automatic gain control (AGC) in hearing aids. British Journal of Audiology, 22, 93-104.

Moore, B. C. J, Glasberg, B. R., and Stone, M. A. (1991). Optimization of a slow-acting automatic gain control system for use in hearing aids. British Journal of Audiology, 25, 171-182.

Moore, B. C. J, Johnson, J. S., Clark, T. M., and Pulvinage, V. (1992). Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss. Ear and Hearing, 13, 349-360.

Moore, B. C. J, Laurence, R. F., and White, D. (1985). Improvements in speech intelligibility in quiet and in noise produced by 2-channel compression hearing aids. British Journal of Audiology, 19, 175-187.

Moore, B. C. J, Lynch, C., and Stone, M. A. (1992). Effects of the fitting parameters of a two-channel compression system on the intelligibility of speech in quiet and in noise. British Journal Audiology, 26, 369-379.

Nabelek, I. V. (1983). Performance of hearing-impaired listeners under various types of amplitude compression. Journal of the Acoustical Society of America, 74, 776-791.

Niemeyer, W. (1965). Speech audiometry with phonetically balanced sentences. International Audiology, 4, 97-101.

Niemoelleer, A. F., McCormick, L., and Miller, J. D. (1974). On the spectrum of the spoken speech. Journal of the Acoustical Society of America, 55, 461.

Owens, E. (1961). Intelligibility of words varying in familiarity. Journal of Speech and Hearing Research, 4, 113-129.

Olsen, W. O., Hawkins, D. B., and Van Tasell, D. J. (1987). Representations of the long-term spectra of speech. Ear and Hearing, 8, 100S-108S.

Olsen, W. O., and Matkin, N. D. (1979). Speech audiometry. In W. F. Rintelmann, Hearing Assessment, Perspectives in Audiology (pp. 133-206). Baltimore: University Park Press.

Olsen, W. O., and Tillman, T. (1968). Hearing aids and sensorineural hearing loss. Annals of Otolaryngology, 77, 717-727.

Pascoe, D. P. (1978). An approach to hearing aid selection. Hearing Instruments, 29, 12-16.

Pavlovic, C.V. (1993). Problems in the predictions of speech recognition performance of normal-hearing and hearing-impaired individuals. In, G. Studebaker and I. Hochberg (Eds.), Acoustical factors affecting hearing aid performance (pp. 221-234). Boston: Allyn and Bacon.

Pearsons, K. S., Bennett, R. L., and Fidell, S. (1977). Speech levels in various environments. Report no. 3281. Cambridge, MA: Bolt, Beranek, and Newman.

Penrod, J. P. (1994). Speech threshold and word recognition/discrimination testing. In, J. Katz (Ed.), Handbook of Clinical Audiology, 4th edition, (pp. 147-164). Baltimore: Williams and Wilkins.

Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. Journal of the Acoustical Society of America, 83, 2322-2327.

Plomp, R. (1989). Reply to 'comments on 'The negative effects of amplitude compression in multichannel hearing aids in the light of the modulation transfer function' '. Journal of the Acoustical Society of America, 86, 428.

Raffin, M. J. M., and Schafer, D. (1980). Application of a probability model based on the binomial distribution of speech-discrimination scores. Journal of Speech and Hearing Research, 23, 570-575.

Raffin, M. J. M., and Thornton, A. (1980). Confidence levels for differences between speech discrimination scores. Journal of Speech and Hearing Research, 23, 5-18.

Resnick, D. M. (1962). Reliability of the twenty-five word phonetically balanced lists. Journal of Auditory Research, 2, 5-12.

Scharf, B. (1978). Loudness. In E. C. Carterette, M. P. Friedman, Handbook of Perception, 4, New York: Academic Press.

Schwartz, D. M., Bess, F. H., and Larson, V. D. (1977). Split half reliability of two word discrimination tests as a function of primary-to-secondary ratio. Journal of Speech and Hearing Disorders, 42, 440-445.

Schwartz, M. F. (1970). Power spectral density measurements of oral and whispered speech. Journal of Speech Hearing Research, 13, 445-446.

Steinberg, J. C., and Gardner, M, B. (1937). The dependence of hearing impairment on sound intensity. Journal of the Acoustical Society of America, 9, 11-23.

Stevens, S. S. (1955). The measurement of loudness. Journal of the Acoustical Society of America, 27, 815-829.

Stevens, S. S. (1957). Concerning the form of the loudness function. Journal of the Acoustical Society of America, 29, 603-606.

Stevens, S. S., Egan, J. P., and Miller, G. A. (1947). Methods of measuring speech spectra. Journal of the Acoustical Society of America, 19, 771-780.

Tarnoczy, T. H. (1956). Determination of the speech spectrum through measurements of superimposed samples. Journal of the Acoustical Society of America, 28, 1270-1275.

Thornton, A., and Raffin, M. J. M. (1978). Speech discrimination scores modeled as a binomial variable. Journal of Speech and Hearing Research, 21, 507-518.

Villchur, E. (1973). Signal processing to improve speech intelligibility in perceptive deafness. Journal of the Acoustical Society of America, 53, 1646-1657.

Villchur, E. (1983). The evaluation of amplitude-compression processing for hearing aids. In, G. A. Studebaker, and F. H. Bess (Eds.), The Vanderbilt Hearing Aid Report. Upper Darby, PA: Monographs in Contemporary Audiology.

Villchur, E. (1989). Comments on the negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. Journal of the Acoustical Society of America, 86, 425-428.

Wallenfels, H. (1967): Hearing aids prescription. Springfield: Charles C. Thomas, publisher.

Walker, G., Byrne, D., and Dillon, H. (1984). The effects of multichannel compression/expansion amplification on the intelligibility of nonsense syllables in noise. Journal of the Acoustical Society of America, 76, 746-757.

Walker, G., and Dillon, H. (1982). Compression in hearing aids: An analysis, a review and some recommendations, Report # 90, Sydney: National Acoustical Laboratories.

Yanick, P. (1976). Effects of signal processing on intelligibility of speech in noise for persons with sensorineural hearing loss. Journal of American Audiological Society, 1, 229-238.