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.

COSTS AND BENEFITS OF PEAK CLIPPING IN AMPLIFICATION FOR PROFOUND HEARING LOSS

by

FRANCIS N. IGLEHART

A dissertation submitted to the Graduate Faculty in the Ph.D. Program 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: 9959189

UMI®

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 the Ph.D. Program in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1-6-2000
Date

Chair of Examining Committee

1-14-2000
Date

Executive Officer

Professor Harry Levitt

Professor Arlene Neuman

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

COSTS AND BENEFITS OF PEAK CLIPPING IN AMPLIFICATION FOR PROFOUND HEARING LOSS

by

Francis N. Iglehart

Advisor: Professor Arthur Boothroyd

Peak clipping, a common form of output limiting in hearing aids, has potential costs (distortion) and benefits (increased gain) to speech-perception ability. This study examined these costs and benefits for subjects with profound hearing loss under three conditions of clipped (by 15 dB) and unclipped speech in quiet. Phoneme-recognition ability and sensation levels were measured for 16 subjects with profound hearing loss and, for comparison, two subjects with moderate hearing loss. First, speech stimuli were clipped to 15 dB below the highest instantaneous peak value and presented at highest comfortable level. The mean phoneme-recognition score for subjects with profound losses was 18.2%. Second, speech stimuli were presented unclipped with the same instantaneous peak level as clipped speech in the first condition. The mean score for the same subjects was 13.1%. Third, unclipped speech was presented at subjects' highest comfortable level. The mean score was 17.1%. Comparisons of these scores show the benefit of increased gain available through the use of clipping and the lack of significant cost of distortion for hearing-aid users with profound losses. The benefit provided by clipping was equal to a 44% increase in channels of independent information (k = 1.44). Sensation levels were also measured. The mean level for subjects with profound hearing loss was 24.1 dB. For the two subjects with moderate hearing loss, the mean sensation level was considerably higher at 41.7 dB. Unlike the subjects with profound hearing loss, the phoneme-recognition scores for the two subjects with moderate loss were relatively unchanged across the three conditions. The relatively large sensation levels for these two subjects permitted a 15 dB attenuation of the signal without loss of phoneme recognition. Under all conditions, vowels in consonant-vowel-consonant words were more easily recognized than either initial or final consonants, and initial consonants were more easily recognized than final consonants.

To Karen, Austen and Katy

CONTENTS

Chapter 4

Chapter 5

APPENDIX

TABLES

Table

FIGURES

Figure

Chapter 1

INTRODUCTION AND PURPOSE

1.1. General purpose

The narrow dynamic range of hearing accompanying a profound loss presents a major challenge to efforts to balance need for gain with need for comfortable listening levels. High levels of gain may benefit the hearing-aid user by providing access to weaker speech sounds yet may, at the same time, amplify stronger sounds beyond the listener's tolerance for loudness if output is not limited properly. Two methods of output limiting commonly used are peak clipping and compression. Each of these approaches to limiting has different effects on the speech signal and, subsequently, the perception of speech by a hearing-aid user. There is, unfortunately, very little information available on the effects of output limiting on speech perception by subjects with profound hearing loss.

The general purpose of the study was to determine the effects of peak clipping on speech-perception performance for listeners with profound hearing loss. This study provided empirical data on the costs and benefits of peak clipping adjusted to permit effective limiting of excessively loud speech sounds while retaining adequate amplification of the weaker speech sounds. The "costs" were the negative effects of peak clipping, or rather, its associated distortions, on speech perception. The "benefits" were the positive effects of increased gain – beyond the point at which clipping begins – on speech perception.

1.2. Definition of terms

For the purposes of this study, all measures of performance are obtained with stimuli presented auditorially without visual cues.

Speech perception refers to the correct identification of language units such as words, syllables, phonemes, and sub-phonemic features such as voicing, manner, and place of articulation.

Phoneme recognition refers to the correct identification of phonemes in lists of consonant-vowel-consonant words.

Phoneme-recognition score is operationally defined as the percentage of phonemes recognized in lists of consonant-vowel-consonant words.

Profound hearing loss refers to a three-frequency-average threshold (at 500, 1000, and 2000 Hz) greater than 90 dBHL (ANSI, 1989).

Highest comfortable level is operationally defined as the highest output of speech stimuli that a subject found comfortable and for which a 2 dB increase in level was reported as too loud or uncomfortable.

Peak clipping refers to a method of amplitude limiting in which the positive and negative excursions of the waveform are not permitted beyond predetermined values.

Intensity range of speech refers to the difference in decibels between the level at which vowel peaks first become audible and the level beyond which further increases add no perceptually useful information. At constant talker effort and distance, this level is considered, for the purposes of this study, to be approximately 30 dB.

Dynamic range of hearing refers to the difference in decibels between threshold of audibility and threshold of discomfort.

Instantaneous peak level is operationally defined as the level in decibels of the largest excursion of the speech waveform as sampled from three word lists used in the present study.

Rms vowel peak is operationally defined as the rms level in decibels of the largest excursion of the portion of the speech waveform representing vowel sounds averaged over a time frame of 100 msec. and sampled from three word lists used in the present study.

1.3. Outline and rationale

1.3.1. Saturation sound pressure level

All hearing aids have a limit beyond which the instantaneous output voltage cannot rise. This limit is commonly referred to as the Saturation Sound Pressure Level (SSPL) and is inherent in the design of electronic amplifiers. Most hearing aids allow for reduction of SSPL below the absolute limit of the amplifier. Selection and adjustment of SSPL is part of the fitting procedure. The goals are that the SSPL should be far enough above the user's threshold of hearing to encompass the intensity variations in the incoming speech signal without distortion, but not so high as to cause discomfort or risk further damage to residual hearing.

1.3.2. SSPL and profound hearing loss

For many hearing aid users, it is difficult to meet both goals for adjustment of SSPL. This problem is especially serious for those with profound hearing loss.

In these subjects, the dynamic range of hearing is often less than that needed to encompass the intensity variations within speech of constant effort produced at a fixed distance. And it is well below the range needed to encompass the additional variations occurring because of changing talker, talker effort and talker distance.

1.3.3. SSPL and peak clipping

Peak clipping is a traditional solution to limiting the output of a hearing aid. Clipping does not permit output sound pressure to rise beyond a certain level. The potential benefit of this approach is that weaker sounds of speech can be raised above the threshold of audibility – potentially enhancing speech-perception performance. The negative consequence is that the stronger sounds of speech are distorted – potentially reducing both speech-perception performance and perceived quality. It is reasonable to assume that when the amount of clipping is small, the benefits of increased audibility might exceed the costs of distortion. As gain and clipping are increased, however, there should come a point at which the costs of distortion begin to negate the benefits of increased audibility. Unfortunately, the relationship between speech-perception performance and the amount of peak clipping in subjects with profound hearing loss has received little attention.

1.3.4. SSPL and compression limiting

In fact, the current trend in hearing aid design is to avoid peak clipping altogether by the use of various forms of compression. In compression, the output signal is prevented from exceeding preset levels by introducing temporary reductions in gain. The reduction in gain may remain in effect after an input signal has passed, at which point the gain then returns to its original value. It is not clear, however, that compression is an appropriate solution to the dynamic range problems of persons with profound hearing loss. The temporary reduction of gain may render important information inaudible during activation of compression. It may also reduce the audibility of weak sounds that occur during the time that gain is returning to its pre-compression value. In other words, compression may have a greater negative effect on speech perception than would have been produced by the distortion the compression is designed to avoid. Once again, however, there are few empirical data on this issue.

1.4. Specific goals

The specific goals of this study were to measure the effects of clipping on intelligibility of speech in quiet for subjects with profound hearing loss: a) when clipped speech is presented at highest comfortable level and unclipped speech is presented at the same instantaneous peak level and b) when both clipped and unclipped speech are presented at highest comfortable level. The research questions were:

1. Is clipped speech more intelligible than unclipped speech when the clipped speech is presented at highest comfortable level and the unclipped speech is presented at the same instantaneous peak level? This question addresses the potential benefit of increasing rms level in the signal without increasing the instantaneous peak level.
2. Is clipped speech more or less intelligible than unclipped speech when both are presented at most comfortable level? This question addresses the

potential costs of peak clipping in terms of reduced intelligibility caused by distortion.

From a clinical perspective, the question is whether it is appropriate to increase gain in a linear hearing aid beyond the point at which clipping begins, thereby increasing intelligibility but also distortion, or should one seek another means, such as compression, for increasing the rms level that does not exceed the listener's tolerance for loud sound.

Chapter 2

LITERATURE REVIEW

2.1. Intensity range of speech and dynamic range of profound hearing loss

2.1.1. Introduction

One challenge to the use of amplification is to fit the wide intensity range of speech into the relatively narrow dynamic range of a profound hearing loss. The range of intensity in the speech signals arriving at a hearing aid is considerable. Even when speech is produced at a constant average level, there are marked moment-to-moment variations above and below that average. These variations occur both across frequency and over time. To these variations must be added the changes of average level that accompany changes of talker, talker effort, and talker distance. Thus the intensity range of speech can be well over 50 dB. The dynamic range of hearing with a profound loss, on the other hand, is usually considerably smaller – on the order of 25 dB or less. The challenge is to provide gain to the speech signal and yet not exceed the listener's loudness tolerance.

2.1.2. Speech intensity at constant effort and distance

Dunn and White (1940) provided one of the earliest measurements of the wide-ranging amplitude of speech. Their measurements help illustrate the difficulty of perceiving a speech signal within the relatively narrow dynamic range of profound hearing loss. Dunn and White demonstrated the high average intensity of speech in the region of the first vocal tract formant (around 200 to 600 Hz) and the decrease in intensity of about 6 dB per octave at frequencies above and below this region. They also showed that, at any single frequency, the short-term intensity varies above and below the average, with a range of 40 or 50 dB. It should be noted, however, that the exact range of variation was dependent on the time window over which speech energy was integrated. It should also be noted that the lower end of the measured intensity range would be determined by the noise floor in Dunn and White's equipment, rather than by the properties of the speech itself. The result may have been an underestimation of the full intensity range of speech. At the same time, Dunn and White did not establish the intensity range of speech signals that contributes to speech intelligibility – a value important to studies of speech perception.

2.1.3. Intensity range of useful information in speech

The range of intensity variation in important speech information at a given frequency is generally assumed to be 30 dB (French and Steinberg, 1947). This is the value incorporated into the original development of the Articulation Index (ANSI, 1969; Kryter, 1962) and, more recently, the Speech Intelligibility Index (ANSI, 1997).

Behavioral evidence for the 30 dB intensity range of speech of constant average level comes from the work of Miller (1947) and Boothroyd (1968). Miller studied normally hearing subjects listening to speech in a fixed background of random noise. Boothroyd studied normally hearing subjects listening in quiet at near-threshold levels. Both showed that when speech has been adjusted so that the vowel peaks are just audible, a further 30 dB increase of average level is needed before recognition of the weaker consonants is attained. In other words, the 30 dB range of the normal Performance-versus-Intensity function provides a measure of the effective intensity range of speech of constant average effort. It should be noted, however, that this behavioral intensity range of speech applies to the complete speech signal. It is somewhat less than would be predicted from a 30 dB intensity range at any frequency combined with an average spectrum that falls at 6 dB per octave above 600 Hz.

Yet another approach to estimating the effective intensity range of speech of constant average level is to measure the intensities of individual speech sounds. Fletcher (1950) showed that the strong vowel /aw/ is some 30 dB more intense than the weak fricative /th/. Similar findings have been reported more recently by Boothroyd, Erickson and Medwetsky (1994). These last authors also showed, however, that there are marked inter-subject differences of dynamic range.

2.1.4. Speech intensity varying by effort, distance or talker

The designers, fitters, and users of hearing aids must also deal with the changes of average speech intensity that accompany changes of talkers, talker effort and talker distance.

The most extensive reference on the effects of changing talker effort comes from Pearsons, Bennet and Fidell (1976). They found that raised speech, loud speech, and shouting are, respectively, 7, 16, and 28 dB more intense than speech of normal effort. Moreover, the shape of the average spectrum changes with vocal effort. Boothroyd (1993a) has suggested a range of 10 dB below the average to encompass the range of soft speech. He also suggests 10 dB above the average as a rough estimate of the intensity of loud speech.

The effects of varying talker distance are dependent on the reverberation times of the listening environment, which are different at different frequencies. In an anechoic environment, the effect of talker distance can be estimated by the inverse square law. In everyday listening environments, the inverse square law works reasonably well up to a few feet – up to the point at which the reverberation sound level exceeds the direct sound level. This "critical distance" depends on the reverberant characteristics of the room and is greater for low-reverberation environments. When expressed in terms of decibels, the inverse square law predicts a change of 6 dB for every doubling (or halving) of the distance between the talker and the microphone of the listener's hearing aid (Boothroyd, 1993a)[1]. If we assume, therefore, that the listener's own speech is produced six inches from the hearing aid microphone while a conversational partner is four feet distant (and within the critical distance of that listening environment), it seems reasonable to predict that the difference in average speech levels will be 18 dB.

From the foregoing, it may be concluded that the intensity range within speech of constant average effort is in the region of 30 dB; that variations of talker effort can change the average level by plus or minus 10 dB; and that the difference of average level between speech at conversational distance and self-

---

[1] Data from Pearsons et al. (1976) indicate that speech levels in classrooms decrease by about 6 dB when measurement distance doubled from one meter to two. Their speech-level data thus seem to follow the inverse square law, at least for conversational distances.

generated speech is on the order of 18 dB. Combining these levels gives a total intensity range at the hearing aid microphone of 68 dB. This is the difference between a strong vowel in the loud speech of the hearing-aid wearer and a weak fricative in the soft speech of a conversational partner. Even if we discount the differences of effort, the range is still in the region of 48 dB. As mentioned earlier, fitting this range of intensities into the dynamic range of hearing of persons with profound hearing loss presents a serious challenge to the designers and fitters of hearing aids (Levitt, 1982).

2.1.5. Dynamic range of hearing in persons with profound losses

The dynamic range of hearing, as defined earlier, refers to the difference in decibels between the threshold of audibility and the threshold of discomfort. In normally hearing individuals, this range is in the region of 100 dB between 500 and 4000 Hz (Dadson and King, 1952; Wegel, 1932). It is a characteristic of sensorineural hearing loss that there is often a relatively unchanged threshold of discomfort in spite of dramatic changes in threshold of audibility (Skinner, 1988, p. 27). The result is a serious reduction of the dynamic range of hearing. Moreover, the dynamic range tends to be different at different frequencies – in inverse relationship to the elevation of audibility threshold. The fitting of hearing aids needs to consider the threshold of audibility and the threshold of discomfort. Too high an SSPL can cause loudness discomfort and additional hearing loss from exposure to excessively loud sound; too low a level can lead to unnecessary saturation and an inadequate sensation level (Dillon and Storey, 1998).

From the point of view of hearing aid fitting, a more practical definition of dynamic range of hearing is the difference, in decibels, between the speech detection threshold, also known as the speech-awareness threshold, and the highest comfortable level. Boothroyd (1993b) has provided information on dynamic range of hearing, measured in this way, in prelingually deafened subjects with losses in the severe to profound range. He found that the average dynamic range fell by roughly 5 dB for every 10 dB increase in hearing loss. The average dynamic range was 30 dB for losses in the region of 80 dBHL. Between 90 dB and 110 dBHL, the average dynamic range fell from 25 dB to 15 dB. It should be noted, however, that these were average values. Individual values of dynamic range varied by +/-15 dB at any given hearing loss. Moreover, for some of the subjects, maximum listening level was determined by the amplification equipment rather than their discomfort levels. Nevertheless, these data demonstrate that even speech of constant average effort cannot fit within the dynamic range of hearing of many persons with profound losses. Even if some persons can attain adequate audibility of conversational speech, the additional 18 dB in the intensity of self-generated speech will drive the hearing aid into limiting. If this limiting involves peak clipping, then distortion will occur and perception of self-generated speech may be affected.

2.2. Segmental speech perception by persons with profound hearing loss

Although discrimination is seriously impaired, persons with profound hearing loss retain considerable ability to use segmental information in the acoustic speech signal – at least for losses up to around 110 dBHL and, for some individuals, even beyond. Subjects with this level of hearing loss typically perceive certain segmental features more accurately than other features. For example, consonant voicing and manner are typically perceived more accurately than is place (Boothroyd, 1984; Flynn, Dowell and Clark, 1998; Merklein, 1981; Smith, 1975). The problems with feature perception carry over to perception of categories. For example, recognition of vowels is typically more accurate than of consonants (Boothroyd, 1984; Flynn et al, 1998; Merklein, 1981; Pickett, Martin, Johnson, Smith, David, Willis and Otis, 1972). Degrees of hearing loss within the range of profound can be used to predict speech-perception performance across groups of subjects. The performances by individuals with a given level of hearing loss, however, differ considerably (e.g., Boothroyd, 1984; Erber and Alencewicz, 1976; Flynn et al, 1998; Pickett et al, 1972).

2.3. Peak clipping and its consequences

2.3.1. Introduction

Peak clipping is a common method in hearing-aid design for output limiting. It also permits the application of additional gain to a signal while protecting the comfort of the listener. It is a relatively simple and inexpensive method of amplitude limiting in which the excursion of the waveform is not permitted beyond a predetermined value. Clipping occurs whenever the instantaneous level of the signal, plus the gain of the hearing aid, exceeds the SSPL of the aid.

2.3.2. Electroacoustic characteristics of peak clipping

As mentioned earlier, the potential benefit of peak clipping is that it can permit a hearing aid, within its limited maximum output, to produce a signal of higher rms value. Up to a point, the greater the degree of clipping allowed, the higher the overall output level. For example, Wathen-Dunn and Lipke (1958) measured the increase in long-term-averaged power when a speech signal was clipped and then returned to the same instantaneous peak level that existed prior to clipping. The authors reported that 24 dB of clipping followed by amplification to the pre-clipping peak level yielded about 12 dB of additional power. Twelve decibels of clipping yielded about 8 dB of additional power.

Clipping may also be used to reduce the intensity range of a signal. As discussed earlier, the intensity range of speech of constant effort and distance is in the region of 30 dB and this can easily exceed the dynamic range of many persons with profound hearing loss. Clipping is one means to address this lack of "fit". For example, if peak clipping is used to reduce the rms peaks of the speech signal by 12 dB then the overall intensity range would be reduced from 30 dB to 18 dB. The reduced intensity range of 18 dB in the clipped signal may then more closely fit the reduced dynamic range of a profound hearing loss.

Peak clipping, however, inevitably causes distortion and the more severe the clipping, the greater the distortion (Licklider, 1946). Distortion is defined here as the presence of output waveforms that are not exact, proportional reproductions of the input waveforms. One component of distortion, harmonic distortion, consists of energy at integral multiples of each tonal component in the input spectrum. Another component, intermodulation distortion, consists of energy at frequencies that are the sums and differences of the various tonal components in the output, including the components of harmonic distortion. In general, harmonic distortion in hearing aids is not very serious for three reasons. First, any harmonics created by the distortion were most likely present in the original waveform anyway. Second, the harmonic energy tends to be in the higher frequencies which are not reproduced well by the hearing aid's output transducer. Third, and perhaps most importantly, the harmonic energy tends to be in frequencies higher than the aidable residual hearing of a profound loss. Intermodulation distortion, however, is very serious for two main reasons. First, when there are two speakers, or background noise, much of the energy in the distortion is at frequencies that were not present in the input. Second, because of the difference components, much of the distortion energy is in the low frequencies (Preves and Newton, 1989) where persons with profound losses typically have better residual hearing.

In summary, clipping can permit an increase in the power of a speech signal while preventing the signal from exceeding the listener's tolerance for loudness. Peak clipping, however, may cause distortion at any point when the amplitude of the waveform exceeds a predetermined limit. The distortion includes harmonic and intermodulation distortion.

2.3.3. Perception of clipped speech signals and normal hearing

The earliest studies of peak clipping and speech-perception performance focused on subjects with normal hearing. The perception of clipped speech depends on many factors – the severity and type of clipping, whether stimuli are presented in quiet or noise, where noise is introduced in the circuit, the type of speech test used, and any frequency shaping introduced.

Licklider (1946) studied the effects of peak clipping on the perception of speech in quiet and in "ambient" noise for an unspecified number of normally hearing subjects. Monosyllabic words were spoken live (Licklider, 1944). The noise was introduced at either the talker's or listener's positions relative to the peak-clipping circuit and at a speech-to-noise (S/N) ratio of +20 dB. The types of peak clipping used were symmetrical, asymmetrical and infinite. He found that the presence of peak clipping, even infinite in degree, had remarkably limited effect on perception of speech in quiet and speech in noise that had been added after processing. Clipping was more detrimental to perception of speech when non-impulsive noise was added prior to processing when compared to speech in quiet or in noise added after processing. It was also shown that speech-perception performance in quiet was minimally reduced by processing with either symmetrical or asymmetrical peak clipping. In fact, clipped speech was more intelligible than unclipped speech when both signals were presented to the listener at the same instantaneous peak level. It is important to note that the clipping advantage occurred only when the instantaneous peak level was at less-than-optimal sensation levels and intelligibility scores for both clipped and unclipped speech were below 100 percent. At higher, more optimal sensation levels the difference in intelligibility scores disappeared between clipped and unclipped speech perception.

Licklider and Pollack (1948) studied the perception of speech (Psycho-Acoustic Laboratory Phonetically-Balanced lists; Egan, 1944) in quiet, processed with infinite clipping at the same rms level as unclipped speech, by five normally hearing subjects. They reported that the clipped speech was only moderately intelligible. However, intelligibility rose to 90 percent with practice. (Licklider, 1946, found far better speech-perception performance with clipping possibly because he was using experienced listeners.) Licklider and Pollack also reported that, with subjects' repeated exposure to clipped speech, clipping had little effect on speech perception when compared to unclipped speech at the same rms level.

Pollack and Pickett (1959) studied the effects of various levels of symmetrical clipping (0 to 24 dB) on speech perception in noise. The symmetrically clipped signals were restored to the original unclipped rms level. The speech materials were Harvard Phonetically-Balanced monosyllabic word lists. The five subjects had normal hearing and were trained in listening tasks. Two types of noise (white and low-frequency) were added after clipping and at two loudness levels. The authors reported that, across all the listening conditions, speech-perception scores were relatively unchanged. That is, the clipped signal, regardless of clipping level, was as intelligible as unclipped speech when presented at the same rms level. The authors noted, however, that noise added prior to clipping could have been detrimental to speech perception.

Ainsworth (1967) investigated the perception of transformed clipped speech by 12 subjects with presumably normal hearing. Phonetically balanced monosyllabic word lists in quiet were processed under 20 transforms of clipping in order to determine the acoustic information within a clipped speech signal that most contributed to intelligibility[1]. The author found that information on polarity of the zero crossings was important to the intelligibility of speech but neither the polarity of the pulses nor polarity of the zero crossings needed to match the original speech waveform. Relevant to the present study, the data from use of these 20 transforms indicated, overall, that subjects confused vowel-like sounds to a lesser extent than fricative-like consonantal sounds. Ainsworth felt that "clipped noise" filled the silent gaps of speech and obscured more fricative-like sounds.

Bode and Kasten (1971) studied the speech-perception performance of 34 normally hearing subjects listening to speech at various levels of harmonic distortion (intermodulation distortion was not mentioned) at the same instantaneous peak level and in white noise. The distortion was likely caused by clipping since the authors described the harmonic distortion as the result of increasing gain in hearing aids through which the speech was processed. The Modified Rhyme Test (House, Williams, Hecker and Kryter, 1965) was recorded undistorted and at four levels of harmonic distortion through a hearing aid. White noise was added to the speech signal at S/N 0 dB prior to clipping. Speech-perception scores declined considerably with the introduction of low levels of distortion but further large increases in distortion did not produce an equivalent

---

[1] These processes transformed the zero crossings of speech waveforms into pulses of either the same or reverse polarity. The purpose was to remove all acoustic information but the timing and direction of the zero crossings.

decrease in scores. For example, with the mean speech-perception score of 76 percent for undistorted speech, the introduction of low levels of distortion (15 percent) caused scores to decline considerably to 53 percent. But with substantially more distortion (35 percent), scores declined only slightly further to 47 percent. Thus, the introduction of relatively low levels of distortion had a noticeable impact on speech-perception scores while additional distortion demonstrated less of a proportional effect. Perception of initial consonants was more resilient to the effects of distortion than perception of final consonants. For example, mean perception scores with 15 percent distortion were 62 percent for initial-consonant position and 43 percent for final position. The authors noted that increased peak clipping of speech in white noise reduced the peaks of speech more than the peaks of white noise and, hence, reduced the S/N ratio.

The studies cited so far that have included noise in their listening conditions have used either white or "ambient" noise. There is a greater likelihood, however, that noise affecting speech perception in everyday-listening situations will be competing speech noise. Gioannini and Franzen (1978) observed the perception of speech in speech noise for 30 subjects with normal hearing. The noise was continuous speech at S/N +12 dB introduced prior to processing. The method used for clipping the speech signal was described only as an adjustment of a hearing aid to a condition of high harmonic distortion as measured by ANSI standards (S3.3-1960). Presumably, this condition produced clipping by raising gain until expected output was beyond the SSPL of the aid and the same instantaneous peak level was maintained. The authors reported that scores dropped significantly when 15-percent distortion was introduced. The authors, thereby, demonstrated the deleterious effects of nonlinear distortion on speech-perception performance in speech noise by normally hearing subjects. The authors also reported on the sensitivity of different speech tests in speech noise to different levels of clipping distortion. One speech-perception test, the Modified Rhyme Hearing Test (Kreul, Nixon, Kryter, Bell, Lang and Schubert, 1968), showed relatively little sensitivity to an increase in distortion from one percent to 15 percent. However, another test, the Phonetically Balanced – 50/Competing Message test (based on work by Egan, 1948, and modified by Gioannini and Franzen, 1978), revealed significantly more sensitivity. Subjects with normal hearing demonstrated a 20-percent decrease in speech-perception scores between the low (less than one percent) and high (15 percent) distortion conditions.

2.3.4. Perception of clipped and filtered speech signals and normal hearing

Several studies have examined the effects of clipping on speech-perception performance for normally hearing subjects when frequency shaping is included. Licklider and Pollack (1948) found that use of high-frequency emphasis prior to infinite clipping improved speech-perception performance in quiet to above 90 percent. Others have found similar results (e.g., Thomas and Niederjohn, 1970). One theory attempting to explain the benefit of high-pass filtering prior to infinite clipping states that high-pass filtering 1) provides greater emphasis to the components of speech important to speech perception and

2) reduces the production of distortion products caused by infinite clipping of lower frequencies that would fall within the range of important, higher frequency speech components (Thomas and Ravindran, 1974).

Licklider and Pollack (1948) studied the perception by five normally hearing subjects of infinitely clipped speech (Psycho-Acoustic Laboratory Phonetically-Balanced lists; Egan, 1944) processed in quiet. They found that infinitely clipped speech was moderately intelligible. Scores, however, improved from 70 to 90 percent with listening practice. Infinitely clipped speech with a high-frequency emphasis (6 dB/octave) introduced prior to clipping was more intelligible – scores improved from 90 to 95 percent with practice – than unfiltered, clipped speech. Infinitely clipped speech with a low -frequency emphasis (6 dB per octave) introduced prior to clipping was minimally intelligible – scores improved from near zero to 25 percent with practice. Neither low- nor high-frequency-emphasis filtering applied after clipping affected speech-perception scores. The authors did not attempt to explain the results but did note the considerable benefit to speech perception provided by the use of high-frequency-emphasis filtering prior to clipping. They also noted the improvement in perception scores that accompanied practice and attributed the improvement, in part, to familiarity with the test words.

Pollack (1952) measured the perception of speech (Psycho-Acoustic Laboratories phonetically balanced word lists) that was processed with either low- or high-frequency-emphasis filtering and then infinitely peak clipped. The unspecified number of subjects had presumably normal hearing. White noise was added, after filtering and clipping, at S/N ratios varying from –20 dB to +10 dB. The results were compared to an earlier study in which speech-perception performance was measured for speech stimuli that was filtered but not clipped (Pollack, 1948). Clipped speech in the later study was amplified to the original instantaneous peak level of the unclipped speech of the earlier study. Pollack reported that speech-perception scores for the clipped condition were better than scores for the unclipped condition when speech-to-noise (S/N) ratios in both conditions dropped below +5 dB. The author felt that these results for perception of unclipped speech were observed because, above S/N +5 dB, the weaker, high-frequency sounds were audible against low background noise and contributed more significantly to speech perception than did the stronger, low-frequency components. However, at low S/N ratios, unclipped, high-frequency components were masked by noise and thus did not contribute to the perception of speech. Speech perception improved with clipping because clipping lowered the peaks of the more intense low-frequency speech sounds relative to the level of the weaker, high-frequency speech sounds. As gain was applied and the signal was raised to the pre-clipped level, the high-frequency components were amplified to a level that was louder than within the unclipped signal and became audible again through the noise. Thus, clipping permitted additional high-frequency gain that resulted in better speech-perception performance. Similar to the findings of speech perception in quiet by Licklider and Pollack (1948), the subjects also experienced better perception of clipped speech in noise when the high frequencies were preserved and the low-frequency components of speech were suppressed. These findings provided further evidence that the reduction of low-frequency components before clipping reduces clipping's negative effects on speech perception, at least for subjects with normal hearing.

Thomas and Niederjohn (1970) studied the perception of infinitely-clipped speech with high-pass filtering for ten subjects with normal hearing. Speech (phonetically balanced word lists; Egan, 1948) was processed through various high-pass filters incorporating different cut-off frequencies and slopes. The speech signal was then infinitely clipped and, lastly, combined with white noise. The authors reported that the perception of speech processed prior to the addition of noise improved with the use of high-pass filtering. The highest speech-perception scores were obtained when filtering with a cut-off frequency of 1100 Hz and a slope of 12 dB/octave. Thomas and Niederjohn, thereby, confirmed the importance of high-pass filtering for improving perception of clipped speech when noise was added after processing.

Thomas and Ravindran (1974) studied ten subjects with normal hearing and the perception of speech (Harvard phonetically-balanced word lists) under various processing conditions. The authors used speech stimuli processed through high-pass filtering with a frequency cutoff of 1100 Hz and a 12 dB/octave slope, combined with white noise, and then infinite clipped. Noise added prior to clipping at S/N ratios of 0, 5 and 10 dB provided more hearing aid-like conditions than those used by Thomas and Niederjohn (1970). Thomas and Ravindran found that filtered/clipped speech was about 10 to 20 percentage points more intelligible than unmodified (i.e., unclipped and unfiltered) speech presented at apparently the same rms level and across the three S/N ratios. They felt that the contribution of filtering to the intelligibility of speech was due to the suppression of low-frequency components that contributed relatively less to speech perception and to the emphasis on higher-frequency components that contributed more. As well, the distortion products caused by clipping fell above this important high-frequency range. The authors mentioned the potential benefits that may come from the application of filtering and clipping in hearing-aid circuits in regard to speech-perception performance and hearing-aid use in noisy environments.

2.3.5. Perception of clipped and filtered speech signals and hearing loss

Additional studies of peak clipping, frequency filtering and speech-perception performance have included persons with various levels of hearing loss. Thomas and Sparks (1971) compared the perception of filtered/clipped speech in quiet to the perception of unprocessed (i.e., unclipped and unfiltered) speech presented at the "same average sound pressure level"[1] and same sensation levels for 16 hearing-impaired subjects. The subjects had conductive, mixed or sensorineural losses with a wide range of three-frequency averages (15 dB to 75 dB ISO). Processed speech was high-pass filtered at 1100 Hz (attenuating the first formant) with a slope of 12 dB/octave and then infinitely

---

[1] The average sound pressure level of the filtered/clipped speech was defined as the rms of the signal at the headphones. The average sound pressure level of the unmodified speech was determined by measuring the peak-to-peak values of the waveforms of spoken words appearing on an oscilloscope and assuming a "peak value" of 14 dB.

peak clipped. This processed condition was already known to be highly intelligible to normally hearing listeners listening in noise in comparison to an unprocessed condition (Thomas and Niederjohn, 1970). Speech stimuli were phonetically-balanced word lists (Egan, 1948). The authors reported that most (13 of 17) subjects listening at sensation levels of 20 dBSL and 30 dBSL had better perception scores with filtered/clipped speech than with unprocessed speech. The perception scores for filtered/clipped speech and for unprocessed speech, on the other hand, were comparable at higher sensation levels of 40 dBSL. The authors, however, recommended against the use of high-pass-filtered/clipped processing for persons with profound hearing loss who have no residual hearing in the high frequencies.

2.3.6. Perception of clipped versus compressed signals and hearing loss

Several studies of subjects with hearing loss focused on the intelligibility of speech processed through compression-limiting circuits in comparison to speech processed through peak-clipping circuits. Hudgins, Marquis, Nichols, Peterson and Ross (1948) reported on the comparison of speech-perception scores of hearing-impaired subjects listening through a selection of compression-limiting aids and peak-clipping aids. The speech stimuli were Phonetically-Balanced monosyllables (Egan, 1944). Six subjects with moderate to moderately severe hearing loss responded to words lists presented in noise and processed through three types of hearing aids. One type of aid was a commercially available peak-clipping model with the option of either flat or high-frequency response and an

SSPL of 128 dB. Another commercially-available aid had "automatic volume control" and was linear in output up to 122 dBSPL. At that level, automatic volume control activated with an input/output ratio of about five-to-one and produced a maximum output of 136 dBSPL. Harmonic distortion was limited to about ten percent at outputs near 130 dBSPL. The third type of aid was a master hearing aid in which compression limiting began at outputs of 117 dBSPL with an input/output ratio of about twelve-to-one; distortion never exceeded ten percent. Attack and release times were not provided, though attack times were described as sufficiently fast to limit output of instantaneous peaks in a speech signal. Output was limited to a maximum of 130 dBSPL. Speech perception was reported as superior with both the automatic-volume-control aids and compression-limiting aids as compared to results with the peak-clipping aids but only at high speech-input levels. The authors did not address the possibility that better speech-perception results obtained with the automatic volume-control aid and the compression-limiting aid at high input levels were the result of both aids having higher maximum output levels than available from the peak-clipping aid. In other words, there remains the possibility that had subjects used three types of aids with comparable maximum output, the subjects would have demonstrated comparable speech-perception scores with each of the three aids. At moderate levels of input, most subjects performed equally well with the three instruments. The authors concluded that reduced distortion produced by the compression-limiting aid made this aid a more desirable choice over a wider range of speech inputs when compared to the peak-clipping aid.

Dreschler (1988) compared the effects of peak clipping and compression limiting in hearing aids on phoneme recognition in quiet for 16 subjects with a variety of hearing losses ranging in degree from mild to severe. The peak clipping aids had an SSPL of about 111 dBSPL. Compression in the compression-limiting aids activated at about 104 dBSPL with an input/output ratio of 10:1 and a maximum output of 114 dBSPL. Attack and release times were 6 and 55 msec., respectively. Subjects demonstrated significantly better phoneme recognition with compression-limiting aids than with peak-clipping aids. Dreschler reported that the low-frequency information important for phoneme identification (i.e., first formant) was perceived better in compression-limited speech than in clipped speech. Plosives, however, were perceived better in peak-clipped speech. Dreschler discussed the possible advantage to speech perception provided by the higher output of the compression aid. He noted, however, that once this output advantage was taken into account, a sizable speech-perception advantage with use of the compression aid still remained.

Dawson, Dillon and Battaglia (1990) compared the perception and quality of speech for subjects with severe and profound hearing loss using peak-clipping and compression-limiting hearing aids. Fourteen subjects were tested, roughly half the group with severe-to-profound hearing losses and the rest with profound losses. Speech-perception tests were selected as suitable for each subject's age and language skills. The noise was competing babble at a variety of S/N levels – each level was apparently dictated by the protocol of the specific test used. Subjects with better intelligibility scores at a certain S/N ratio were given a higher "speech-recognition-competence score" than those with poorer intelligibility scores at the same, or a better, S/N ratio. The authors investigated relationships between a number of independent variables and the group speech-intelligibility scores by type of aid. Four variables provided the strongest prediction of better speech-perception performance with peak-clipping aids as compared to compression-limiting aids. These variables were a high degree of saturation of the peak-clipping aids at an Leq input of 70 dBSPL[1], a high preferred volume setting for the peak-clipping aids, low speech-recognition-competence scores, and a high degree of hearing loss.

2.3.7. Peak clipping and quality of speech, loudness-discomfort levels, and speech-reception thresholds and normal hearing

A broader view of peak clipping includes consideration of other aspects of hearing-aid use than just speech intelligibility. Some of these aspects reported in the literature are subjective judgments of speech quality, loudness-discomfort levels and speech-reception thresholds. While subjective judgments of peak clipping were not addressed in the present study, it is important to note that perceived quality can strongly influence the user's acceptance, or proper use, of a hearing aid. Judgments of unacceptable quality may lead to the rejection or improper use of an aid which itself would negate the potential benefit of the hearing aid for purposes of speech perception. The literature shows that clipping

---

[1] It is assumed that the saturation was largely due to the result of high levels of gain. Approximately half the subjects wore their aids set at maximum, or near-maximum, gain and SSPL settings. Complete information on the settings used by other subjects was not provided.

has generally been found to be far more detrimental to speech quality than to speech-perception performance.

Some of the earliest evidence on the effects of peak clipping on listeners' judgments of quality involved subjects with normal hearing. Licklider (1946) noted that, while clipping did not greatly affect scores for speech perception in quiet for subjects with normal hearing, subjects reported that quality was severely affected. Noise was more detrimental to the quality of clipped speech when added to the speech signal prior to clipping compared to the quality of a signal to which the noise had been added after clipping. This finding was to be expected since noise added after clipping does not contribute to intermodulation distortion as does noise added prior to clipping. Licklider's observations of the effects of the entry point of noise into a circuit processing speech is important in the study of speech-perception performance and hearing-aid use. Noise, in an everyday listening environment, is processed with a speech signal, not added afterwards. Thus, a study of the perception of clipped speech in background noise should introduce noise into the speech signal prior to clipping in order to realistically simulate everyday listening conditions.

Licklider and Pollack (1948) also noted that subjects with normal hearing found infinitely clipped speech in quiet, while moderately intelligible, to be poor in quality. The introduction of high-frequency-emphasis filtering (6 dB/octave across the entire frequency range), however, restored quality to near pre-clipping levels. Low-frequency-emphasis filtering (6 dB/octave), on the other hand, produced judgments reflecting extremely poor quality.

2.3.8. Peak clipping and quality of speech, loudness-discomfort levels, and speech-reception thresholds and hearing loss

Several studies of subjective judgments of quality, loudness-discomfort levels, and speech-reception thresholds and peak clipping of speech in quiet have included subjects with hearing loss. The hearing losses were typically moderate to moderately severe in degree.

Lawson and Chial (1982) studied the quality judgments by 12 normally hearing and 12 hearing-impaired listeners using filtered and/or clipped speech (textbook passages) in quiet. The hearing-impaired subjects had various levels of sensorineural hearing loss – generally mild in the lower test frequencies and sloping, by the authors' definition, to moderate-to-profound in the higher frequencies. (Note that the hearing loss of any one of these subjects would not be categorized as profound in degree according to the definition used in this study.) Frequency modification was accomplished by various low-, high-, and band-pass filters. The hearing-impaired subjects were more sensitive to high-pass filtering and distortion than they were to low-pass filtering and, overall, were less sensitive to distortion than were the subjects with normal hearing. The authors felt that subjects with hearing loss may have reduced sensitivity to distortion that had, in this case, spread to the higher frequencies. They felt that speech-quality estimates were a useful method for assessing hearing aids and for the identification of perceptual differences among normally-hearing and hearing-impaired groups. A factor potentially influencing subjective ratings by hearing-impaired subjects was prior experience with hearing aids. An experienced hearing-aid user may favor an experimental listening condition whose sound quality is similar to his or her personal amplification and therefore familiar. Such prior experience with hearing-aid use may influence subjective ratings obtained under experimental conditions. Information on the subjects' personal amplification was not provided in the Lawson and Chial study.

Kozma-Spytek, Kates and Revoile (1996) studied the quality ratings given by a group of hearing-impaired subjects to peak-clipped and high-pass-filtered speech (sentences; Bench, Kowal and Bamford, 1979) in quiet. Of eight subjects, seven had moderate-to-severe losses and one had a profound loss. The subjects, like eight normally hearing subjects of a previous study (Kates and Kozma-Spytek, 1994), could readily detect peak-clipping distortion in speech. These subjects with hearing loss, while disliking the quality of clipped speech in comparison to unclipped speech, were less bothered by it than the normally hearing subjects. The hearing-impaired subjects also gave lower quality ratings to a high-frequency-boost condition than to the low-frequency boost condition. A comparison of relative scores across listening conditions (high-frequency boost, low-frequency boost and flat-frequency response) indicated that these subjects apparently based judgments of speech quality on the distortion perceived at low frequencies (i.e., below 1000 Hz) while ignoring the high-frequency components. The subjects also showed greater variability in the individual rating scales than the previous subjects with normal hearing. This finding indicated to the authors that subjects with different hearing losses use different listening strategies which, in turn, highlighted the risks of applying conclusions based on normally hearing subjects to those with hearing loss. Though the authors did not discuss their findings for the one subject with profound hearing loss, their data indicated that this subject gave the highest quality ratings to speech regardless of clipping level (zero to 12 dB) when the speech was amplified with low-frequency emphasis or a flat-frequency response. As discussed in the previous review, prior experience with personal hearing aids may influence subjective ratings of experimental amplification. Information on the subjects' personal amplification, however, was not provided.

Fortune and Preves (1992) reported on subjects of generally moderate hearing loss (type of loss not provided) and their aided-loudness-discomfort levels and quality ratings for continuous discourse (Speech Intelligibility Rating test; Cox and McDaniel, 1984) in quiet. The authors reported that the five subjects' aided-loudness-discomfort levels and quality ratings had an inverse relationship to the level of distortion of speech. As coherence values[1] decreased from 1.0 to 0.78, the loudness-discomfort levels decreased by a mean of seven decibels. Fortune and Preves noted, therefore, that loudness discomfort may be produced not only by high output levels but also by concurrent distortion. When a loud signal as common as the hearing-aid user's own voice produces not only saturation but also high levels of distortion, the hearing-aid user may be inclined to reduce the aid's volume setting. This new setting may then be too low for general listening benefit.

---

[1] Coherence reflects the extent to which the inputs and outputs are equal in all regards except overall amplitude. The coherence function assigns a value between 0 and 1, with 1 denoting an accurate representation of the input signal and 0, a complete deterioration of the signal (Preves and Newton, 1989).

Crain and Van Tasell (1994) studied the shifts in speech-reception thresholds of four hearing-impaired and ten normally hearing subjects while listening to speech in quiet (Central Institute for the Deaf W-1 spondee list) processed with various levels of peak clipping. The hearing-impaired subjects had either mild or moderately severe losses (type of loss not provided). Both the normally hearing and hearing-impaired subjects showed progressively higher speech-reception thresholds in response to increasing levels of peak clipping and the upward shifts were greater for the hearing-impaired group. An increase in level of clipping (to a level between 18 dB and 24 dB) coincided with both significant increases in speech-reception thresholds and subjective judgments that quality had become unacceptable. However, the shifts in speech-reception thresholds for both groups were small (0.5 dB to 5 dB).

2.4. Summary

In summary, literature on the effects of peak clipping on speech perception covers a wide variety of subjects and test conditions. Subjects were described as having hearing sensitivity that ranged for normal to very impaired and having listening experience with clipped speech that ranged from no prior experience to extensive exposure. These subjects listened to clipped speech stimuli under various clipping levels, in quiet or in different types of noise introduced at different points of the processing circuit, or combined with filtering that employed various slopes and cutoff frequencies.

If clipping alone is considered, two factors are known to play a large role in its effects on speech perception. One is the increase in long-term average power available with clipping when the signal is amplified to its pre-clipped peak level. In other words, clipping permits a system of limited maximum output to produce a signal of higher rms value. Peak clipping, on the other hand, can impair speech perception, especially in the presence of noise, through the introduction of harmonic and intermodulation distortion that increases with the severity of the clipping. These two factors, increased power and distortion, compete as benefits and costs, respectively, in speech perception with hearing aids.

One approach to investigating these costs and benefits for the listener with profound hearing loss is to examine two key issues that emerge in review of the literature. First, clipping benefits speech perception by subjects with normal hearing and subjects with severe-or-better hearing loss – if two conditions are met. These two conditions are i) that the clipped and unclipped speech are presented at the same instantaneous peak levels and ii) that less-than-optimal sensations levels are used. Prior to the present study, data has not been collected under either of these conditions for subjects with profound hearing loss.

Another issue is reflected in the finding that subjects with normal hearing demonstrate the same level of speech-perception ability regardless of whether speech is clipped or unclipped provided that both signals are presented at the same rms level. Some data collected under these conditions for subjects with hearing loss have been published but the effects of clipping are difficult to isolate from other independent variables used in the studies. For example, studies using clipped speech as a listening condition simultaneously employed frequency filtering in the processing strategy. Consequently, there is a need for further data on the effects of clipping on perception by subjects with profound hearing loss when the speech is presented at the same rms level as an unclipped signal. These data also need to be collected, at least initially, without the influence of other variables.

As stated earlier, the present study addressed several of these issues. Speech stimuli were presented without filtering or background noise to subjects with profound hearing loss. Phoneme-recognition scores were first compared under clipped and unclipped conditions at the same instantaneous peak level for the purpose of examining the potential benefit of peak clipping. Phoneme-recognition scores were then compared under clipped and unclipped conditions at the highest comfortable level. This second comparison addressed the potential costs of peak clipping in terms of reduced intelligibility caused by distortion.

Chapter 3

METHODS

3.1. Test conditions

The test methods and conditions used in this study addressed important questions about the effects of peak clipping as a commonly used form of output limiting in amplification for persons with profound hearing loss and the effects of that clipping on speech perception in quiet. These questions, previously unanswered in the literature, were: 1) Is there any benefit to phoneme recognition by profoundly-impaired listeners in amplifying speech signals beyond the level of clipping? This question is addressed by comparison of results from listening conditions #1 and #2 presented to subjects in this study and described below. 2) When peak-clipped speech is adjusted to the listener's highest-comfortable listening level, does the clipping distortion outweigh the benefits provided by increased gain? This question is addressed by comparison of results from conditions #1 and #3 below.

These three listening conditions were (Figure 3.1):

Condition #1: Speech stimuli were clipped to 15 dB below the instantaneous peak value and presented at highest comfortable level.

Condition #2: Speech stimuli were presented unclipped with the same instantaneous peak value as after clipping in the first condition.

Condition #3: Unclipped speech was presented at highest comfortable level for unclipped speech.

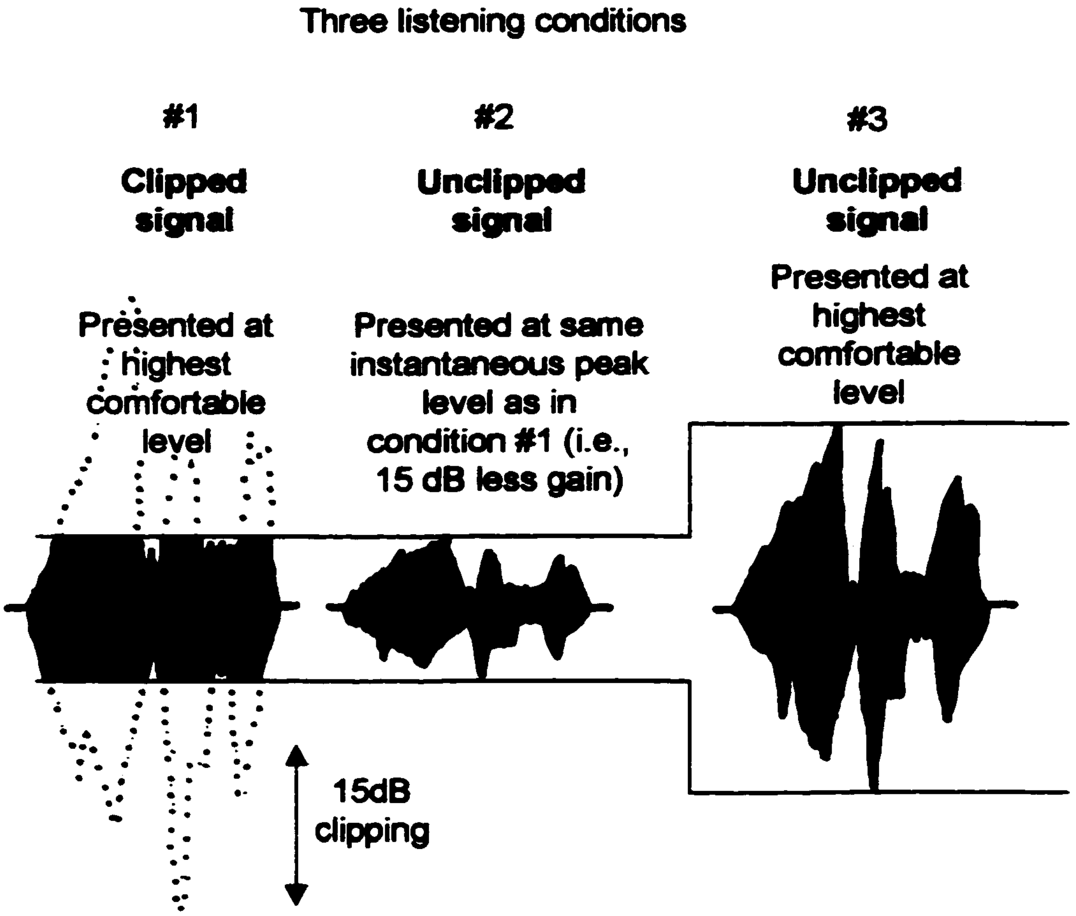


Figure 3.1. Three listening conditions used in the present study. The three waveforms are of the same segment of speech as processed within each of the three conditions.

3.2. Subjects

Subjects were 8 boys and 8 girls with profound sensorineural hearing loss who were enrolled in an auditory/oral educational program. Subjects were selected if they were regular users of hearing aids, possessed good language skills and could attend to auditory-only listening tasks for extended periods of time. Subjects' ages were from 14 to 16 years. Hearing losses averaged 101 dBHL (range of 93 to 108 dBHL). Because of availability, two subjects with moderate sensorineural hearing losses were also included. These two subjects' hearing losses averaged 48 and 60 dBHL. The results for these two subjects will be reported but not included in the statistical analyses. Subjects were asked to participate and both parental and subject written consents were obtained (Appendix A). Pure tone thresholds were obtained for all subjects. Immittance measurements were also obtained to rule out any middle ear involvement. Audiometric data and information on ages, gender and ethnicities are provided in Table 3.1.

3.3. Independent variable

The independent variable was listening condition, represented by three levels. In condition #1, speech stimuli were clipped to 15 dB below the instantaneous peak value and presented at highest comfortable level. In condition #2, speech stimuli were presented unclipped with the same highest instantaneous peak value as in condition #1. In condition #3, unclipped speech

Table 3.1. Summary of subjects' ages, gender, ethnicities, and audiometric data.

					Puretone thresholds of test ear (dBHL)					
Subj	Age (yrs.)	Gender	Ethnicity	Test ear	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	3-freq. ave.
s1	16.5	M	white	R	60	55	50	40	20	48
s2	15.9	M	hispanic	R	55	55	55	70	50	60
s3	14.1	M	white	L	75	75	95	110	110	93
s4	15.0	M	white	R	90	95	95	95	110	95
s5	14.9	F	white	R	80	85	100	110	120	98
s6	14.8	M	white	R	80	85	100	110	115	98
s7	14.3	M	white	L	55	90	100	105	100	98
s8	15.8	F	hispanic	R	90	90	100	105	110	98
s9	15.5	M	white	R	65	90	100	110	115	100
s10	15.0	M	pac. island.	R	90	85	105	110	100	100
s11	16.0	F	white	L	80	90	100	115	105	102
s12	16.5	F	asian	R	100	95	105	105	95	102
s13	14.1	F	afro. amer.	R	85	90	100	115	110	102
s14	16.8	M	white	R	90	95	110	110	115	105
s15	16.5	F	white	L	85	95	110	110	120	105
s16	15.0	F	hispanic	R	80	95	105	115	115	105
s17	15.2	F	white	L	90	95	105	115	120	105
s18	16.2	M	white	L	90	105	110	110	115	108

was presented at highest comfortable level. For condition #1, clipping of the signal was produced off line before the experiment. Individual presentation level was determined at the time of testing. The procedure used for clipping the signal will be described later.

3.4. Dependent variables

The dependent variable was the number of correctly recognized phonemes in lists of consonant-vowel-consonant words. Highest comfortable levels and speech-awareness levels were also measured. Sensation levels for each subject were calculated for clipped speech stimuli and for unclipped speech stimuli. Sensation level was defined as the difference between presentation level and awareness threshold.

3.5. Test materials

Open-set speech perception was measured using recorded isophonemic lists of CVC words (Boothroyd, 1984; Appendix B). These lists contain the most commonly occurring vowels and consonants in American English CVC words. Forty words were presented under each condition. Twelve lists – four lists in each of three test conditions – were used in each test session and these lists were drawn randomly from a pool of twenty word lists. The test recordings were made by a female talker who presented each word at the end of the carrier phrase "The word is ... ."

3.6. Stimulus preparation

The speech stimuli were recorded on a digital audio tape recorder, Panasonic model SV255, by a female speaker of American English from the northeastern United States. Recordings were prepared in an IAC sound-attenuating enclosure with a microphone distance of 18 inches. The analog output of the digital audio tape recorder was re-digitized at 22050 samples per second with a resolution of 16 bits. The digitized files were then edited to provide 100 msec of silence before and after word offset. High-pass filtering of approximately 10 dB/octave below 250 Hz was also used to remove the low-frequency background noise present in the recording environment. All editing was accomplished with DaDisp, a signal-processing software package from DSP Corporation of Cambridge, Massachusetts.

For the purposes of the present study, the sound files were also scaled so that each test word would have the same instantaneous peak amplitude. Each word was imported into a DaDisp worksheet and the peak amplitude was measured. This peak always occurred within the vowel nucleus. The waveform was then multiplied or divided by the amount needed to adjust the peak amplitude to 20,000 digital units. The sound files thus created provided the stimuli for the two "unclipped" conditions of the present study.

To create the sound files for the "clipped" condition, each file was, again, imported into a DaDisp worksheet and amplified by 10^ (15/20) where 15 dB was the desired amount of clipping[1]. The resulting waveform was then symmetrically clipped so as to have a maximum instantaneous amplitude of plus or minus 20,000 digital units.

---

[1] Assuming the intensity range of speech is 30 dB, 15 dB of clipping was chosen because it would likely have a significant effect on audibility without introducing an excessive amount of distortion. This issue will be explored in more depth in the discussion.

3.7. Stimulus presentation

Stimuli were presented from digitized recordings on a laptop computer to Stimulation Presentation Systems and headphones. Speech stimuli were presented in quiet only and amplified by Stimulus Presentation Systems that had broad-band, flat frequency-response characteristics (Appendix C). Each Stimulus Presentation System included a preamplifier, a power amplifier, and output attenuation for two channels that were separately adjustable in 2 dB steps up to a maximum undistorted rms value for pure tones of 136 dBSPL. Amplification levels within a test session depended on the settings of the computer output control, the preamplifier gain settings of the Stimulus Presentation Systems, and the settings of the output attenuators of the Stimulus Presentation Systems. Only the settings of the output attenuators were changed in the course of the study. TDH50 earphones with supra-aural cushions were used with the Stimulus Presentation Systems. Five Stimulation Presentation Systems were available for testing up to five subjects at one time.

3.8. Test environment

Testing was conducted in a quiet classroom setting with an ambient noise level of approximately 40 dB(A). The quietest listening levels for subjects were during speech-awareness testing and subjects with the best pure tone thresholds had the lowest listening levels. The lowest listening levels for these subjects were at least 30 dB above the ambient noise level of the test room. A background level of noise of 40 dB(A), therefore, was sufficiently low so as not to affect test results.

3.9. Design

The study used a single-factor, three-level within-subjects design in order to measure the main effect of listening condition.

The three levels of listening condition were presented in the same sequence throughout the study. It was necessary to use the same sequence for several reasons. The subjects' schedules limited the amount of time available for completion of each test session. In order to have completed the tasks within the allotted time, it was necessary to minimize both redundancy of tasks and subjects' confusion. The tasks performed within each session were therefore in a nonrandomized sequence that most efficiently utilized the available time. While randomization of the three levels of listening conditions would have been more desirable in this design, a randomized sequence would have taken too long to complete within a session. There was another drawback to randomizing the order of listening conditions. First, it was necessary to start each session with subjects listening to speech at comfortably loud levels in order to establish highest comfortable levels. Then, in situations where condition #2 would be presented first, gain would be reduced. This reduction would result in subjects performing phoneme-recognition tasks at presentation levels well below their preferred listening levels. It was predicted (and, in fact, confirmed during testing) that subjects would object to being tested at low presentation levels. It was thought to be undesirable to begin test sessions at such low levels.

Three to five subjects were tested simultaneously.

3.10. Procedure

3.10.1. Stimulus Presentation Systems.

Each Stimulus Presentation System (SPS) was calibrated for frequency accuracy, harmonic distortion, and accuracy of sound pressure level in both right and left channels before the experiment began and after it concluded. For each calibration, harmonic distortion was 2.5% or less for frequencies 125 Hz through 8000 Hz with one exception – distortion in one channel (left channel, SPS #12) was measured at 2.6% which may have been the result of limitations in high-frequency measurements intrinsic to acoustic couplers (ANSI, 1996, subclause 6.1.5). Accuracy of sound pressure level was within 3 dB for 125 through 3000 Hz, within 4 dB for 4000 Hz, within 6 dB at 6000 Hz and within 15 dB at 8000 Hz. The inaccuracy of sound pressure levels at 4000 Hz and above was not considered significant to the present study due to the lack of auditory access to those frequencies by the subjects with profound hearing loss. Calibration was performed with a Larsen Davis Laboratories calibration system. Each earphone was coupled to a 6cc coupler with microphone model # 2575 and output measured by sound-level meter model # 800B and calibrator model CA250.

A listening check of the channels was performed immediately prior to, and following, each test session. These daily checks involved the experimenter listening through each headphone to the calibration tone for consistent loudness between headphones and Stimulus Presentation Systems, a check for unwanted acoustic signals and a physical manipulation of cords and plugs to verify circuit continuity. No Stimulus Presentation System or headset was suspected of malfunction.

3.10.2. Instructions, selection of test ear, highest comfortable levels

The subjects were seated in a semi-circle in a quiet classroom setting and they were not able to easily see each other's test sheets (Figure 3.2). At the beginning of a test session, each subject was required to read written instructions that described the task (Appendix D)[1]. Each subject then chose either the right or left ear for testing. The majority of subjects clearly preferred the opportunity to choose what they believed to be their "better ear" for speech perception. If the subject had no preference, the decision was based on the ear with the better three-frequency-average threshold. The same ear was used for the two test sessions. After placement of the headphones, the subjects were asked individually to judge his/her highest comfortable level for the clipped stimuli. The tester, in a method-of-limits procedure, presented speech stimuli clipped by 15 dB and increased the output level in 2 dB steps to the level judged by the subject as "too loud". This was done one subject at a time with the help of an assistant. The speech stimuli were the carrier phrase ending in words from a list not used in the speech-perception portion of the test session. The test listening level was adjusted to 2 dB below the lowest level judged by the subject as "too loud" and to

---

[1] Note that these written instructions included the request that subjects adjust their attenuators ("volume") themselves while listening to speech sounds. This request was then verbally modified such that the research assistant adjusted the attenuator knobs and the subject verbally reported when loudest comfortable level had been reached. This instruction was changed at the beginning of the first session. The modified instructions and procedure were used throughout the study.

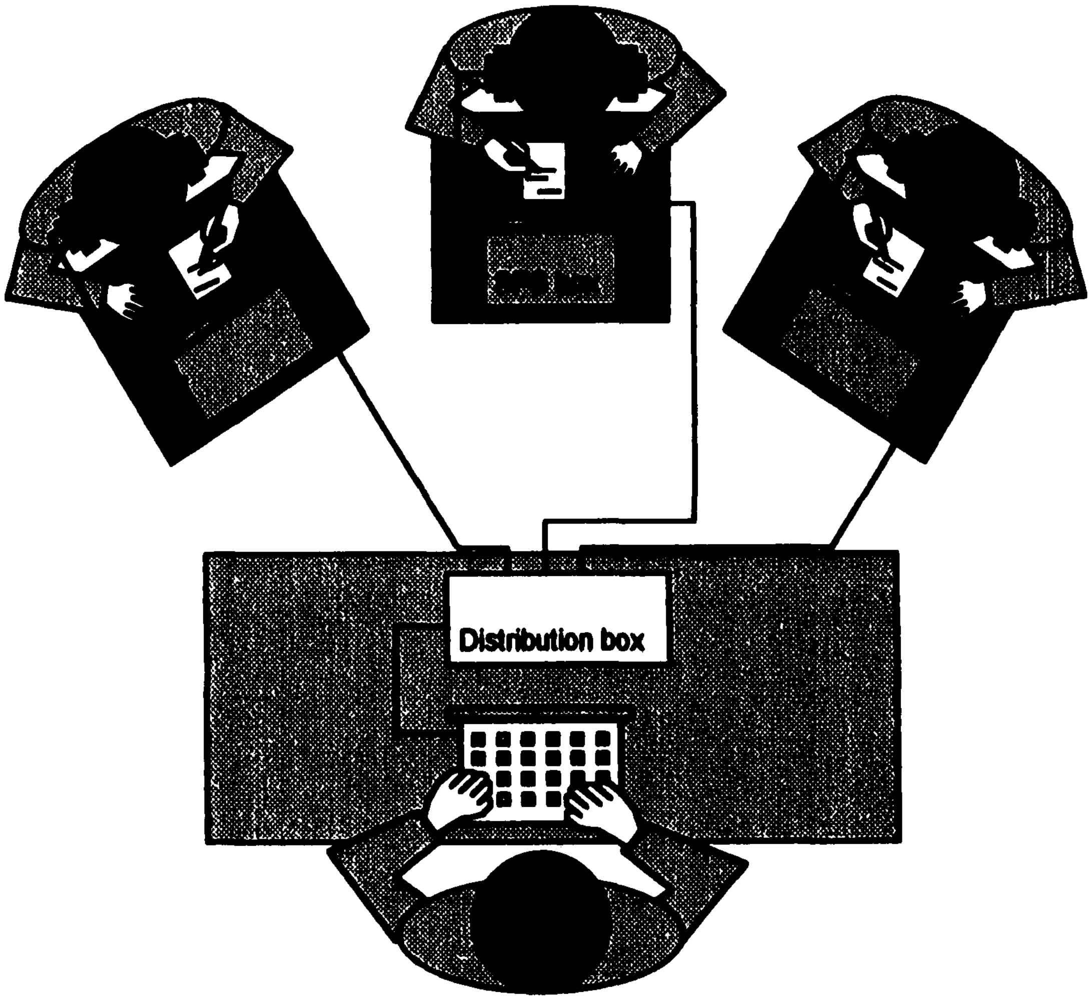


Figure 3.2. Experimental arrangement for speech perception testing.

a level judged as "comfortable". This measurement was repeated until the subject reported the same highest comfortable level at least two consecutive times. The tester exercised caution in order to avoid inadvertently encouraging a subject to tolerate uncomfortable loudness even for brief moments during the test session. Once the testing was underway, no subject reported that the output of the headphones was too loud.

Care was taken to reassure the subjects that no one was judged on performance and that guessing was an acceptable response. A potential problem is the tendency for subjects, when tested in the area in which they have a profound disability, to become discouraged when they sense that performances are especially poor. Condition #2 was very difficult for these subjects.

Each test session lasted about 45 minutes. The type of testing and its length of time were very similar to routine audiological testing that was already familiar to the subjects and to the tester. Neither fatigue nor boredom, therefore, appeared to be a significant problem.

3.10.3. Phoneme-recognition testing

After the highest comfortable level was established for each subject under condition #1, the subjects were instructed again not to change their volume controls. An assistant watched the subjects to ensure that this instruction was followed. Subjects were asked to write down their responses and listening condition #1 began. Condition #1 concluded after 40 words had been presented. Listening condition #2 began with instructions to the subjects that the next words might sound soft, but that they were not to change their volume controls.

Condition #2 began and forty words were presented. The subjects wrote down their responses, and then condition #2 concluded. At the onset of condition #3, the volume controls were re-adjusted to highest comfortable levels while subjects listened to unclipped speech stimuli. The stimuli were, again, the carrier phrase ending in words from a list not used in the speech-perception portion of the test session. Once the highest comfortable level for unclipped speech had been measured for each subject, a third set of forty words was presented. Subjects wrote down their responses. Condition #3 and the entire speech-perception portion of the test session were now complete. By this point, highest comfortable levels had been calculated separately both for clipped speech stimuli and for unclipped speech stimuli.

3.10.4. Speech-awareness thresholds

Measurements of speech-awareness thresholds were then obtained. The same procedure was used in establishing speech-awareness thresholds as was used in measuring highest comfortable levels, but with the following change. The method of limits procedure concluded on an ascending 2 dB step above the highest level at which each subject reported no awareness of speech and at the level at which awareness of speech began. This measurement was repeated until the subject reported the same speech-awareness threshold at least two consecutive times.

A speech-awareness threshold was measured for each of two listening conditions. First, this threshold was measured as the subject listened to a carrier phrase and word list clipped by 15 dB (condition #1). Second, this threshold was measured as the subject listened to a carrier phrase and word list that were not clipped (conditions #2 and #3). At this point, the test session concluded. Sensation levels for each subject were later calculated by subtracting each speech-awareness threshold from the respective highest comfortable listening level.

3.10.5. Repetition of test session

The session was repeated one additional time within a few weeks. Word lists were newly selected by randomization. Highest comfortable levels and speech-awareness levels were again measured. A monetary "thank-you" gift was given to each subject at the conclusion of the study.

Chapter 4

RESULTS

4.1. Phoneme-recognition data

Phoneme and word recognition scores will be found in Appendix E. Provided are the percent recognition of initial consonants, vowels, final consonants and whole words. Scores are shown separately for the two sessions. The Appendix provides data for all 18 subjects. Means, standard deviations and standard errors, however, are shown only for the 16 subjects with profound hearing loss. This practice will be followed throughout this chapter. That is, the data for all 18 subjects will be reported and illustrated but all statistical treatments will apply only to the subjects with profound hearing loss[1].

4.2. Analysis of variance in the phoneme-recognition data

The phoneme-recognition scores for subjects with profound hearing loss were subjected to three-way analysis of variance. Factors were Session (at two levels), Listening Condition (at three levels), and Phoneme Type (at three levels). The results are shown in Table 4.1. It will be seen from this Table that the main effects of Listening Condition and Phoneme Type were highly significant. The main effect of Session, however, failed to reach the five percent level of significance, as did all interactions. Of importance is the absence of significant interaction between Listening Condition and Phoneme Type. From this finding, we can conclude that there is no evidence that the effect of listening condition,

---

[1] The focus of the study was on the 16 subjects with profound hearing loss. The data for the two subjects with moderate losses are reported, however, for the additional insights they may provide.

Table 4.1. Summary of analysis of variance for all effects of Session (S), Listening Condition (C), and Phoneme Type (T). Analysis is based only on data from subjects with profound hearing loss.

Effect	df effect	MS effect	df error	MS error	F	p-level
S	1	221.38	15	128.60	1.72	0.2092
C	2	730.53	30	86.85	8.41	0.0013
T	2	4102.02	30	58.62	69.98	< 0.0001
S x C	2	47.29	30	56.38	0.84	0.4422
S x T	2	13.43	30	26.70	0.50	0.6096
C x T	4	30.31	60	31.28	0.97	0.4313
S x C x T	4	23.23	60	25.04	0.93	0.4539

when measured in terms of percent phonemes correct, differs for the three phonemes types. Note that this analysis was performed on percent-correct scores. When the data were arcsine transformed, to increase homogeneity of variance, and the analysis was repeated, the findings were exactly the same[1]. The only significant effects were the main effects of Listening Condition and Phoneme Type.

4.3. Main effect of Listening Condition

The group mean phoneme-recognition scores for the three listening conditions (collapsed across phoneme type and session) were as follows:

Condition #1: Clipped speech at Highest Comfortable Level –	18.2%
Condition #2: Unclipped speech with the same instantaneous peak level –	13.1%
Condition #3: Unclipped speech at Highest Comfortable Level –	17.1%

In post hoc testing based on the analysis in Table 4.1 (using the least significant difference test) it was found that the mean score for condition #2 was significantly different from condition #1 (p = 0.0006) and condition #3 (p = 0.0031). Means for conditions #1 and #3, however, were not significantly different from each other (p = 0.5529). From these findings, it may be concluded that, for subjects with profound hearing loss, 15 dB clipped speech is more intelligible than unclipped speech when the two are presented at the same instantaneous peak level. There is no evidence, however, that the distortion introduced by 15 dB of clipping has a significant effect on intelligibility when clipped and unclipped speech are both

---

[1] As the focus of the study was percent phonemes correct – not transforms of percent correct – the use of arcsine-transformed data was limited in these analyses.

presented at Highest Comfortable Level.

The average phoneme-recognition scores for the two subjects with moderate hearing losses under condition #1 was 88.5%, under condition #2 was 89.5% and under condition #3 was 91%. Unlike the subjects with profound hearing loss, there was no evidence of a drop in scores between conditions #1 and #2.

Table 4.2 shows the mean phoneme and word scores of the 18 subjects under the three listening conditions. Data are collapsed across phoneme type and session. Also shown are the means, standard deviations and standard errors for the 16 profoundly deaf subjects. The relationship between scores obtained under conditions #1 and #2 is illustrated in Figure 4.1. The curve shows the least-squares fit to the equation:

$$y = 100^{*}(1 - (1 - x/100)^{k}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where

y = percent phoneme recognition under condition #1

x = percent phoneme recognition under condition #2

and

k = an exponent representing the effective proportional increase in the channels of independent information made available by raising gain 15 dB beyond the point at which clipping first begins.

Table 4.2. Phoneme and word recognition scores (in percent) for three listening conditions. Scores are collapsed across replication. Phoneme scores are collapsed across initial, medial and final positions.

Subj	Hearing loss 3-freq. ave. (dB)	Listening condition (scores in %)					
		#1		#2		#3	
		Phonemes	Words	Phonemes	Words	Phonemes	Words
s1	48	93	84	96	90	94	86
s2	60	84	69	83	61	88	74
s3	93	19	0	16	1	14	0
s4	95	40	15	17	0	26	8
s5	98	16	1	13	0	19	1
s6	98	3	0	9	0	9	0
s7	98	28	10	23	3	27	5
s8	98	14	3	10	1	10	0
s9	100	34	6	26	6	43	16
s10	100	18	1	9	0	14	1
s11	102	23	4	18	1	23	5
s12	102	10	0	7	1	10	3
s13	102	27	8	23	6	25	3
s14	105	14	0	11	1	10	0
s15	105	17	1	7	0	20	3
s16	105	7	0	5	0	9	0
s17	105	12	0	6	0	10	0
s18	108	10	1	9	0	10	0
Means for profound		18.2	3.1	13.1	1.3	17.4	2.7
sd		9.9	4.4	6.6	2.1	9.6	4.3
se		2.5	1.1	1.7	0.5	2.4	1.1

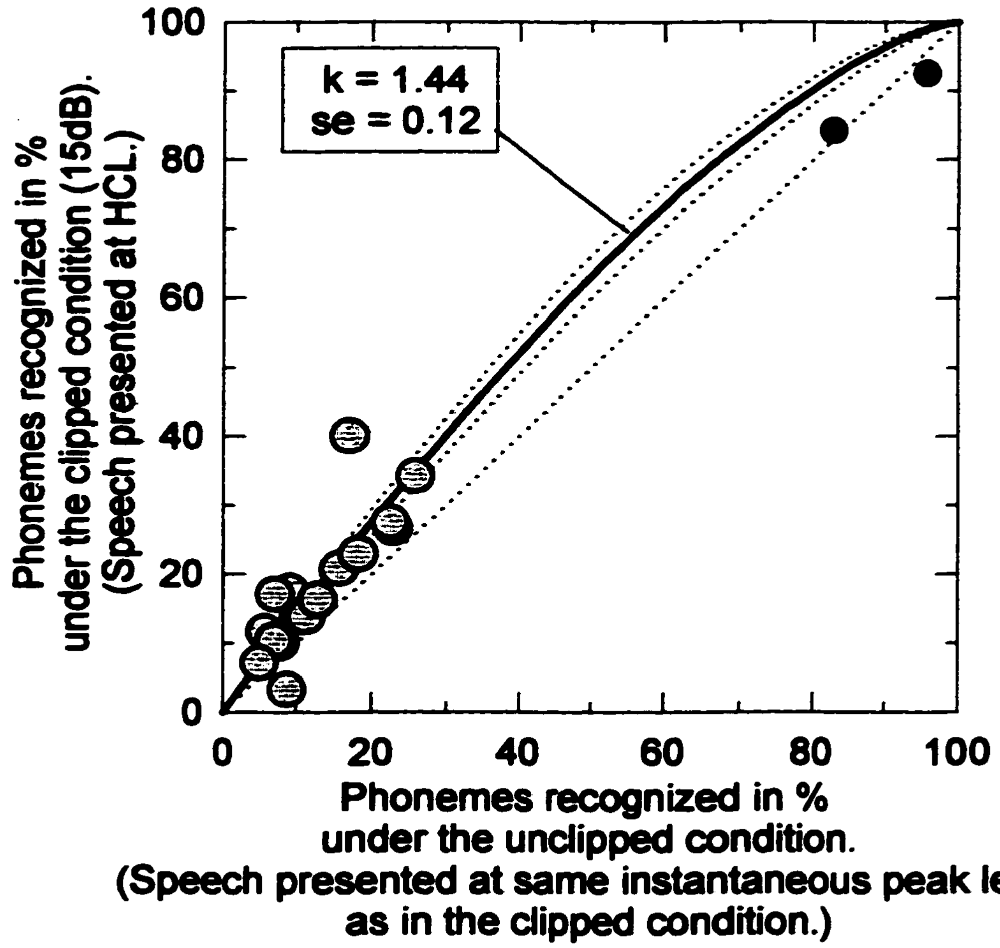


Figure 4.1. Comparison of phoneme recognition scores for clipped and unclipped speech when the two have the same instantaneous peak level - and the clipped speech is presented at highest comfortable level. Curve shows least-squares fit (+/- 1 se) of Equation 1 to the data for 16 subjects with profound hearing loss (large circles).

This equation is based on the work by Boothroyd and Nittrouer (1988) on the effects of context. It has also been used by Boothroyd and Iglehart (1998) to quantify the benefits of FM amplification.

Note that the value of k = 1.44 (se = 0.12) in Figure 4.1 was based only on the data for the 16 subjects with profound hearing loss. This value of k indicates that, for this group of subjects, the benefit of 15 dB of clipping, by allowing for 15 dB of additional gain, was equivalent, on average, to multiplying by 1.44 the channels of independent information available to the listeners, i.e., a 44% increase.

The relationship between scores obtained under conditions #1 and #3 is illustrated in Figure 4.2. As in the previous Figure the data are fit to equation (1) using a least squares criterion. The resulting value of k is 1.04 with a standard error of 0.08. This k value is not significantly different from the value of 1.0, which would represent no difference between the two conditions. Thus, as indicated earlier, this finding provides no evidence to support the conclusion that clipped and unclipped speech are, on average, significantly different in intelligibility to profoundly deaf persons when both are presented in quiet at Highest Comfortable Level.

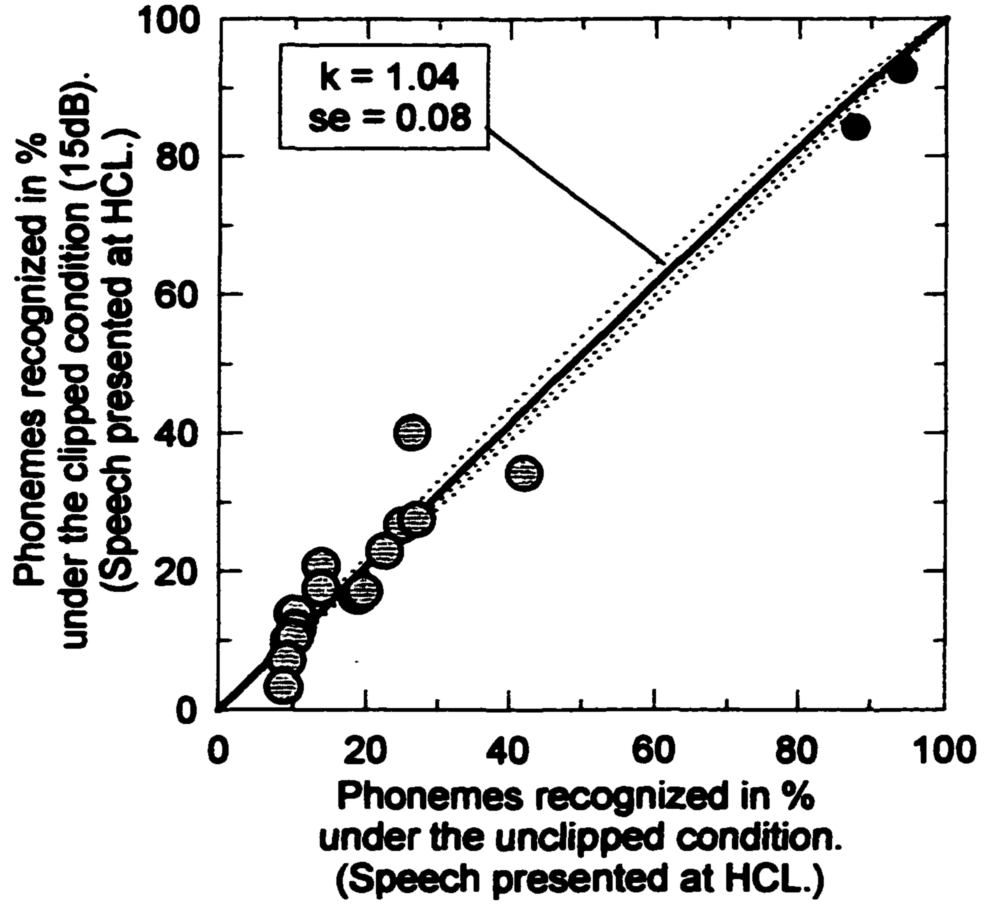


Figure 4.2. Comparison of phoneme recognition scores for clipped and unclipped speech when the two are presented at highest comfortable level. Curve shows least-squares fit (+/- 1 se) of Equation 1 to the data for 16 subjects with profound hearing loss (large circles).

4.4. Main effect of phoneme type

The group mean phoneme-recognition scores for the three phoneme types (collapsed across listening condition and session) were as follows:

Initial consonant – 14.1%
Vowel – 23.6%
Final consonant – 11.0%

These data are illustrated in Figure 4.3.

In post hoc testing based on the analysis in Table 4.1 (using the least significant difference test) it was found that the mean for each phoneme type was significantly different from the other two at least at the 0.01 level. Thus, it may be concluded that, for the average subject with profound hearing loss listening to consonant- vowel-consonant words under the conditions of this experiment, vowels are more easily recognized than are initial or final consonants and initial consonants are more easily recognized than are final consonants.

4.5. Interaction between Listening Condition and Phoneme Type

The analysis of Table 4.1 gives no evidence of an interaction between Listening Condition and Phoneme Type. The conclusions just reached about differences in recognition of phonemes in consonant-vowel-consonant words must, therefore, be assumed to apply to each of the three listening conditions. The actual mean recognition scores for the three conditions and three phoneme types are shown in Table 4.3.

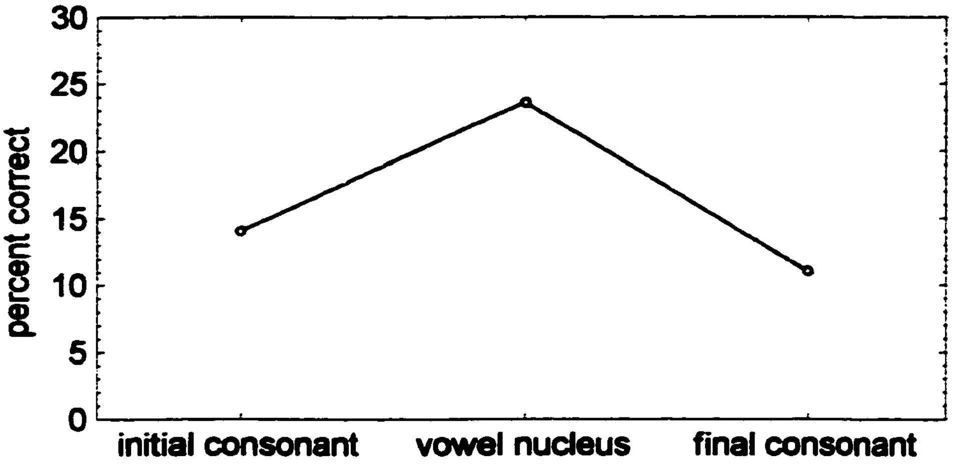


Figure 4.3. Group mean phoneme recognition scores for the three phoneme positions, obtained for 16 subjects with profound hearing loss. Scores are collapsed across three listening conditions.

Table 4.3. Group mean percent phoneme-recognition scores for three phoneme types, obtained by 16 subjects with profound hearing loss listening under three conditions. Scores are collapsed across session.

	Initial consonant	Vowel nucleus	Final consonant
Condition #1	16.7	25.9	12.0
Condition #2	10.8	19.6	8.8
Condition #3	14.7	25.2	12.3

These data are illustrated in Figure 4.4.

4.6. Listening levels

The attenuator readings for individual settings of Speech Awareness Threshold (SAT) and Highest Comfortable Level (HCL) are provided in Appendix F. Also shown are Sensation Levels obtained by subtracting SAT from HCL. Data for condition #1 (HCL and SAT measured with clipped stimuli) and condition #3 (HCL and SAT measured with unclipped stimuli) are provided separately and are shown for each session and for the average of the two sessions. Data are presented for all 18 subjects but means, standard deviations and standard errors are based only on the 16 subjects with profound hearing loss.

The mean sensation level for the 16 subjects with profound hearing loss for the clipped condition (#1) was 25.8 dB with a standard deviation of 5.5 dB. The mean sensation level for the unclipped condition (#3) was 22.4 dB with a standard deviation of 6.0 dB. The overall mean sensation level for these 16 subjects was 24.1 dB. Because condition #3 involved 7 dB more gain than

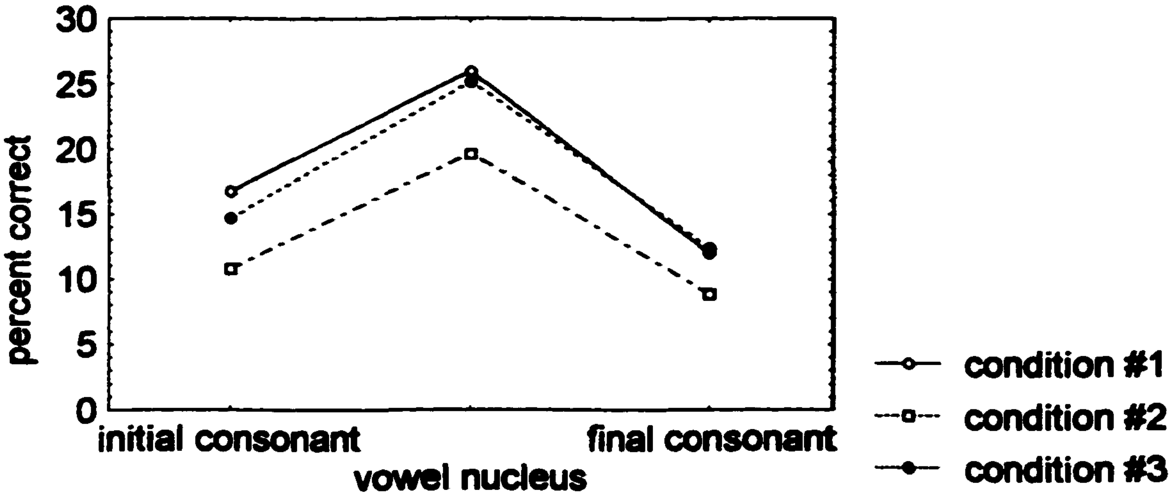


Figure 4.4. Group mean phoneme recognition scores for the three phoneme positions and three listening conditions, obtained by 16 subjects with profound hearing loss.

condition #2, the average sensation level for condition #2 may be assumed to be 15.4 dB.

The mean sensation levels for the two subjects with moderate hearing loss for the clipped condition (#1) was 43.0 dB and for the unclipped condition (#3) was 40.5 dB. The overall mean for these two subjects was 41.7 dB. In comparison to the subjects with profound hearing loss, these mean sensation levels were greater by 17.6 dB. Because condition #3 for these two subjects involved 8 dB more gain than condition #2, the average sensation level for condition #2 may be assumed to be 32.5 dB.

4.7. Analysis of variance in listening levels

The SAT and HCL data for the 16 subjects with profound hearing loss were subjected to a three-way analysis of variance. The factors were Listening Condition at two levels (clipped and unclipped), Listening Measure at two levels (SAT and HCL), and Session at two levels. The result is shown in Table 4.4. It will be seen that session failed to reach the 0.05 level of significance, either as a main effect or in interaction with other effects. This finding supports the collapsing of data across the two sessions. The significant interaction between Condition and Measure implies that the sensation level was different for test conditions #1 and #3.

4.8. Mean listening level data

Table 4.5 shows the group mean attenuator settings at Speech Awareness Threshold and Highest Comfortable Level for clipped and unclipped

Table 4.4. Analysis of variance in the data for Session (S), Listening Condition (C), and Listening Measure (attenuator settings at Speech-Awareness Threshold and Highest Comfortable Level) (M) for 16 subjects with profound hearing loss.

Effect	df effect	MS effect	df error	MS error	F	p-level
S	1	1.53	15	16.06	0.10	0.7618
C	1	2397.781	15	6.85	350.15	< 0.0001
M	1	18576.28	15	58.68	316.56	< 0.0001
S x C	1	11.28	15	10.61	1.06	0.3189
S x M	1	57.78	15	18.45	3.13	0.0971
C x M	1	87.78	15	7.25	12.11	0.0034
S x C x M	1	3.78	15	8.71	0.43	0.5201

speech, together with Sensation Level (HCL – SAT). Note that the attenuator settings are just that. They do not represent actual levels of the speech signal. The sensation levels, on the other hand, are a true measure of the difference between the speech level at which the listener decided that the speech was just audible and that at which the speech was presented for testing under conditions #1 and #3. As indicated earlier, the difference of 3.4 dB between the Sensation Levels for conditions #1 and #3 was statistically significant.

Table 4.5. Group mean attenuator settings for 16 subjects with profound hearing loss listening to clipped (by 15 dB) and unclipped speech expressed in terms of attenuator settings.

	Clipped (15 dB)	Unclipped
Speech Awareness Threshold (dBSPL)	99.4	109.8
Highest Comfortable Level (dBSPL)	125.2	132.2
Sensation Level (dB)	25.8	22.4

To convert the attenuator settings of Table 4.5 into measures of speech amplitude, it was necessary to examine the amplitude distribution in the clipped and unclipped speech. This procedure is described in Appendix G. Several measures were made of speech level – instantaneous peak, rms peak (over 100 msec), long-term rms, average of the rms (over 100 msec) peaks of the vowels, modal rms (100 msec) level, lowest rms peaks of the vowels (over 100 msec) and rms of the 20th percentile of the distribution. The measure of average of rms peaks of the vowels was very similar in level between the clipped and unclipped conditions. The results are shown in Table 4.6. It will be seen that the settings chosen for listening in conditions #1 and #3 were, on average, within 1.0 dB of each other and in the region of 123.2 dBSPL. This correspondence in chosen levels for the two conditions suggests that, when subjects were asked to adjust speech levels to highest comfortable level, they adjusted the levels so that the average rms (100msec) of the vowel peaks was the same for each condition[1]. In condition #2, the listening level was, on average, 7.0 dB lower[2] – at around 116.2 dBSPL.

---

[1] Initially, subjects were presented with backward-playing speech and asked to judge their HCLs. The objective was to use speech stimuli of exactly the same intensity as the target speech while not giving subjects prior exposure to actual test words to be used. When perception testing subsequently began with forward-playing speech, a large portion of subjects complained that the speech was too low in volume. In other words, many subjects reported backward-playing speech as louder than the same speech played forward. As a result, HCLs were then re-established with forward-playing speech. This procedure resulted, for many subjects, in higher attenuator settings. Perception testing was subsequently conducted at HCLs based on forward-playing speech. Some of the subjects also volunteered descriptions of the backward-playing speech such as "It sounds funny."

[2] This was the amount by which, on average, the subjects with profound hearing loss increased the attenuator setting when adjusting the unclipped speech back to HCL between conditions #2 and #3.

Table 4.6. Mean listening levels for 16 subjects with profound hearing loss listening to clipped (by 15 dB) and unclipped speech expressed in terms of the average of the rms vowel peaks in word lists. Subjects were listening at highest comfortable level.

	Clipped (15 dB)	Unclipped
Speech Awareness Threshold (dBSPL)	97.9	100.3
Highest Comfortable Level (dBSPL)	123.7	122.7
Sensation Level (dB)	25.8	22.4

4.9. Repeatability of listening levels

The group mean settings for SAT and HCL were quite repeatable. It became clear during the experiment, however, that individual subjects varied considerably between the two sessions. To obtain estimates of this variability, the data for SAT, HCL, and SL were examined separately in analyses of variance. Each analysis examined two factors – Listening Condition and Session. The final error term in each analysis was used to provide an estimate of the standard deviation of repeated measurements for each variable within subjects – excluding the effects of listener, listening condition, and the main effect of session. The results are shown in Table 4.7. They indicate that a single measure, as performed in this study, can estimate a profoundly deaf subject's "true" listening level with 95 percent confidence to within +/- 6.8 dB for Speech Awareness Threshold, +/- 6.3 dB for Highest Comfortable Level, and +/- 8.8 dB for Sensation Level. Although these numbers seem large, they are in keeping with other reports in the literature – as will be explored in the discussion.

Table 4.7. Estimates of confidence limits of a single measure of listening level based on analysis of the variance in the data for 16 subjects with profound hearing loss.

Measure	Error variance (dB^2)	Standard deviation (dB)	95% confidence limits (dB)
Speech Awareness Level	10.5	3.2	+/- 6.8
Highest Comfortable Level	8.9	3.0	+/- 6.3
Sensation Level	17.4	4.2	+/- 8.8

4.10. Predictors of phoneme recognition

For each of the three test conditions, pure tone thresholds and the three levels of listening condition were entered as potential predictors in a stepwise multiple regression analysis for predicting phoneme-recognition scores (Table 4.2). These analyses were limited to the subjects with profound hearing loss.

For phoneme recognition when listening to clipped speech at Highest Comfortable Level, the best single predictor was threshold at 1000 Hz, accounting for 28 percent of the variance ($r = 0.53$, $p = 0.03$). The addition of thresholds at 500 and 250 Hz gave a multiple correlation of $R = 0.67$.

For phoneme recognition when listening to unclipped speech at the same instantaneous peak level as the clipped speech, the best single predictor was threshold at 250 Hz, accounting for 42 percent of the variance (r = 0.65, p = 0.008). The addition of threshold at 1000 Hz and the Speech Awareness Threshold for unclipped speech gave a multiple correlation of R = 0.79.

For phoneme recognition when listening to unclipped speech at Highest Comfortable Level, the best single predictor was threshold at 250 Hz, accounting for 37 percent of the variance (r = 0.61, p = 0.01). The addition of threshold at 2000 Hz gave a multiple correlation of R = 0.64.

There is negligible evidence from these findings to suggest that level of attenuator settings and sensation level account for significant amounts of variance in individual phoneme-recognition performance beyond that accounted for by pure tone threshold. Degree of hearing loss can, however, account for significant amounts of variance – especially hearing loss in the lower frequencies.

4.11. Phoneme recognition as a function of listening level

By combining data from Section 4.3 and Table 4.6, it is possible to plot group mean phoneme recognition, for subjects with profound hearing loss, as a function of speech level – expressed in terms of the average peak level of vowels in the test words. The result is shown in Figure 4.5. It will be seen that mean phoneme-recognition scores improve with higher rms levels of the vowel peaks.

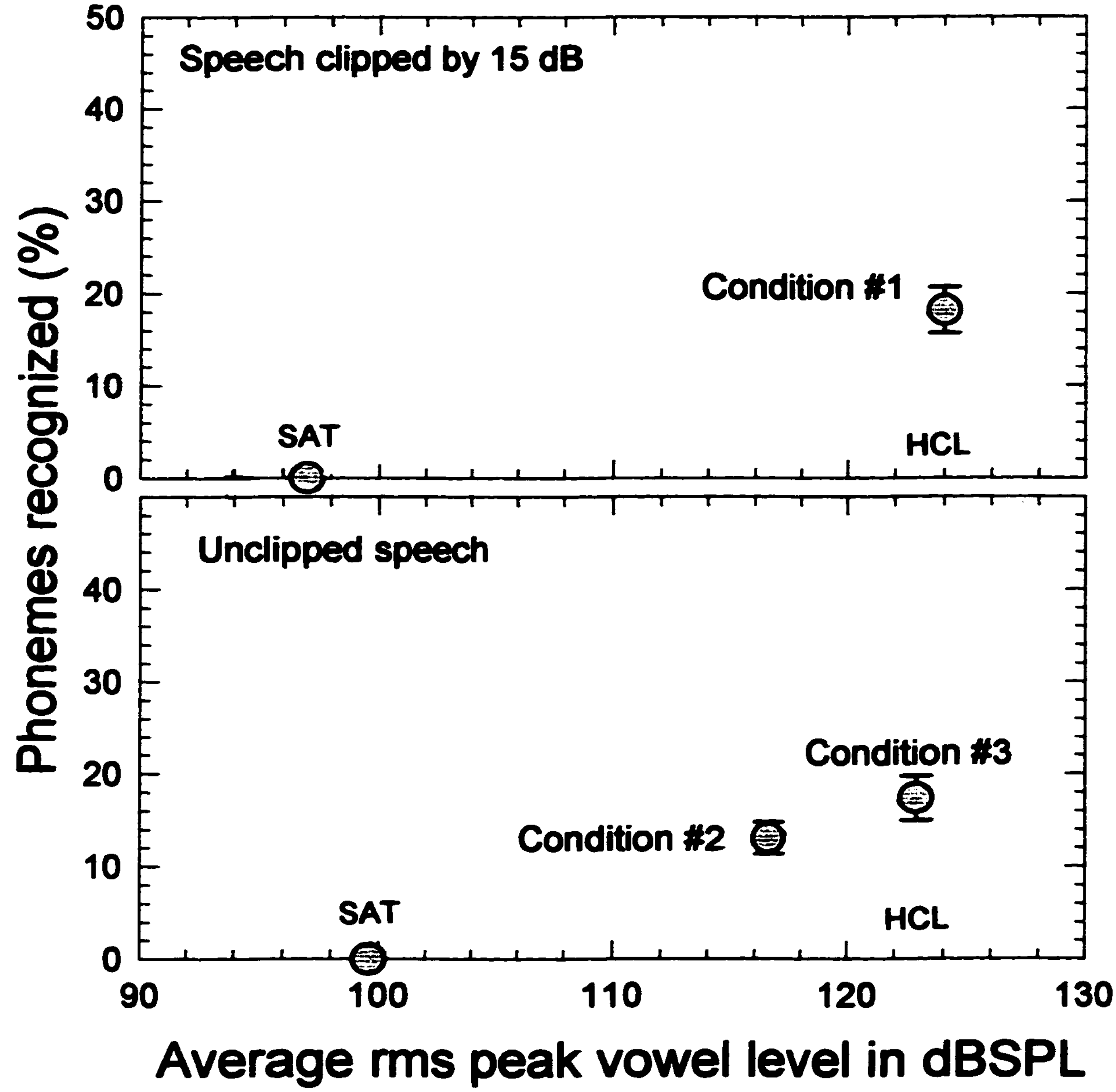


Figure 4.5. Mean phoneme recognition (+/- 1 se) for 16 profoundly deaf subjects as a function of speech level (average of the peak rms levels of vowels) for clipped and unclipped speech. A score of 0% is assumed at Speech-Awareness Threshold.

4.12. Summary

Clipped speech was significantly more intelligible than unclipped speech presented at the same instantaneous peak level. In contrast, there was no significant difference in the intelligibility of clipped and unclipped speech when both were presented at highest comfortable level. The factor of phoneme type was highly significant in consonant-vowel-consonant words with vowels more easily recognized than either initial or final consonants and initial consonants more easily recognized than final consonants across all three listening conditions.

Mean listening levels for clipped and unclipped speech at highest comfortable level were nearly equal when expressed in terms of the average of the rms vowel peaks. Group mean speech-awareness thresholds, expressed in the same terms, were also very similar. Across sessions, however, subjects varied considerably in choice of settings for speech-awareness thresholds and highest comfortable levels.

Chapter 5

Discussion

5.1. Summary of findings

The two research questions proposed earlier concerning subjects with profound hearing loss were answered in this study. First, clipped speech was found to be more intelligible to this group of subjects with profound hearing loss than unclipped speech when the two conditions were presented at the same instantaneous peak level. Second, clipped speech was neither significantly more nor less intelligible than unclipped speech when both were at the highest comfortable level. Other data showed that matching for highest comfortable level equated the two signals in terms of rms level of the typical vowel peaks. Under all conditions for consonant-vowel-consonant words, vowels were more easily recognized than either initial or final consonants, and initial consonants were more easily recognized than final consonants.

For the two subjects with moderate hearing loss, the lack of significant difference in phoneme-recognition scores between the three listening conditions may be explained by the subjects' relatively high, average sensation levels of 42 dB (Appendix F). With these high sensation levels, the reduction of signal gain in condition #2, compared to the other two conditions, apparently did not reduce enough of the speech signal to below these subjects' speech-awareness thresholds to significantly affect phoneme-recognition scores. The distributions of rms amplitudes in Appendix G suggest that an insignificant portion of the speech signal under condition #2 became inaudible to the subjects with moderate hearing loss compared to conditions #1 and #3.

5.2. Previous literature

Similar findings are reported in the literature either for normally hearing subjects or for hearing-impaired subjects. No previous study has, however, specifically targeted subjects with profound hearing loss using a phoneme-recognition task and flat-response, broadband amplification.

A notable issue in this study is that the benefits of clipping can be demonstrated at the phonemic level in a population with profound hearing loss and, therefore, very limited speech-perception capacity. That is, clipping has been shown to have a significantly positive effect on phoneme-recognition ability even in subjects with poor speech-perception ability when clipping permits additional gain in speech signals that do not exceed the listener's comfortable levels.

The finding that, for subjects with profound hearing loss, clipped speech is more intelligible than unclipped speech when the two conditions are presented at the same instantaneous peak level in quiet is consistent with Licklider's (1946) results for subjects with normal hearing. Bode and Kasten (1971) and Gioannini and Franzen (1978), on the other hand, found that clipping had a significantly negative effect on speech perception by subjects with normal hearing. Both of these studies introduced noise prior to processing of the speech signal, but the signal-to-noise level was poorer in the Gioannini and Franzen study. While background noise was not used in the present study, intermodulation distortion caused by noise processed with clipping is an important issue for future research with subjects with profound hearing loss. The effects of intermodulation distortion on speech perception by this population are not well known, yet background noise is common in listening environments.

The improvement in phoneme-recognition scores between the clipped and unclipped conditions at the same instantaneous peak level, as mentioned earlier, was 5.1%. Given that the speech-perception performance of this group of subjects was so poor, a five percentage-point improvement was considerable. As discussed earlier, this improvement in scores is equivalent to a 44% increase in channels of independent information (Boothroyd and Nittrouer, 1988) available to the listeners. From another perspective, the improvement in speech-perception ability provided by clipping, as compared to the unclipped condition at the same instantaneous peak level, may be considered analogous to an improvement in hearing sensitivity. In an earlier study, Boothroyd (1984) reported that one group of profoundly hearing impaired subjects had phoneme-recognition scores that were, on average, better by 18 percentage points than another group of subjects that had 15 dB poorer hearing thresholds (re: three-frequency averages). If the same relationship between speech-perception ability and threshold levels can be assumed to apply to the subjects in the present study, the benefit to speech perception provided by clipping (5.1% improvement in scores), would be comparable to a 4 dB improvement in hearing. In other words, these subjects listening to clipped speech demonstrated speech-perception ability that could be comparable to profoundly deaf subjects with an average of 4 dB better hearing listening to unclipped speech at the same instantaneous peak level.

As noted earlier, the intelligibility of clipped speech did not differ significantly from that of unclipped speech when both were at the highest comfortable level. This finding can be compared to earlier reports that measured perception of clipped and unclipped speech presented at equal rms levels. Comparisons to earlier studies rest on the assumption that subjects in the present study based their highest comfortable level judgements on the rms levels of vowel peaks in speech (Section 4.8). Subjects did, in fact, choose highest comfortable levels with near-equal rms vowel peaks for listening to clipped and unclipped speech (about one decibel difference; Table 4.6). With normally hearing subjects, Licklider and Pollack (1948) and Pollack and Pickett (1959) found no loss of intelligibility with the introduction of clipping (12 dB and 24 dB) when the speech signal was restored to its unclipped rms level. With subjects with hearing loss no worse than severe in degree, Thomas and Sparks (1971) report the same results for listening at high sensation levels. The speech signal in the Thomas and Sparks study, however, was also filtered which makes direct comparison with the present study difficult. For subjects with normal hearing, Thomas and Ravindran (1974), on the other hand, reported better perception of clipped speech compared to perception of unclipped words at apparently the same rms peak level. Comparison to the present study is again difficult because noise had been added after clipping and the signal was high-pass filtered. These issues of noise and filtering will be discussed later.

The group mean recognition scores for the three phoneme types (Section 4.4) show that vowels in the consonant-vowel-consonant words are recognized more easily than either initial or final consonants. Average recognition of vowels in consonant-vowel-consonant words was about 1.9 times more accurate than average recognition of initial and final consonants. It should be noted, of course, that the present study did not include words with vowels and consonants distributed evenly among the initial, medial and final positions. Vowels occurred only in consonant-vowel-consonant words and certain consonants, such as glides, did not occur in final positions. The degree to which equal dispersion of vowels and consonants within word positions could affect test results needs to be explored. Better perception of vowels as compared to consonants, nevertheless, is to be expected largely based on their acoustical properties. Vowels have higher intensity, longer duration and more energy in the lower frequencies where listeners with profound losses typically have their better residual hearing. Consonants, on the other hand, tend to be weaker, short-lived, and to have most of their energy in the higher frequencies.

The proportional advantage in perception of vowels noted in the present study is not inconsistent with several earlier reports. Boothroyd (1984; Experiment 1), for example, found that group mean perception scores for feature contrasts of vowels in consonant-vowel-consonant words by subjects with profound hearing loss were about 1.3 times the scores for perception of feature contrasts of consonants in both initial and final positions. Flynn et al. (1998) found that subjects with severe-to-profound hearing loss (81 to 100 dBHL PTA) recognized vowels, on average, about 1.7 times more accurately than consonants. Flynn et al. had pooled their data on vowel and consonant recognition from the results of several speech-perception tests including tests with vowels spoken in consonant-vowel-consonant words only, consonants spoken in vowel-consonant-vowel words only, and sentence tests. Merklein (1981) reported that subjects with residual hearing in the profound range could recognize features for vowels about 1.6 times better than features for consonants. Merklein based these results on a short list of 18 words comprised of 16 consonant-vowel-consonant words and two consonant-vowel words. Pickett et al. (1972) found that subjects with borderline severe to profound hearing loss could recognize vowels about 1.8 times better than consonants. While Pickett et al. refer to their speech stimuli only as a special version of the Fairbanks-House Modified Rhyme Test, the authors implied that the stimuli were consonant-vowel-consonant words. Though these studies differed in their choices of speech-perception tests, their data agree with the present study that vowels are perceived more accurately than consonants.

In the present study, clipping permitted an increase in presentation level of weaker sounds (i.e., consonants) relative to the subjects' mean threshold (Figure G.2). Clipping tends to equalize the levels of speech sounds (Pollack, 1952). When speech is clipped and pre-clipping peak levels are then restored, the end result is more gain applied to weaker consonants than stronger vowels. The recognition of consonants in the present study, however, did not significantly change between clipped (#1) and unclipped (#3) conditions at highest comfortable level. This finding may be explained by the distributions of rms amplitudes found in Figure G.2. These distributions indicate that weaker speech sounds (i.e., consonants) remained near, or below, the mean threshold level in the clipped condition. For example, weak speech sounds (e.g., weak fricatives such as /f/) were 30 dB below the loudest sounds (vowel peaks) in condition #3 yet were raised in levels only slightly above threshold under condition #1.

It is possible that a higher level of clipping would permit more gain to be applied to the signal and an increase in intensity of the weaker speech sounds. The result, however, may also be increased costs of distortion and no net gain in speech perception. Further work is needed in this area.

Subjects may also have shown better perception of vowels than consonants because of their auditory familiarity with low-frequency speech information. The subjects were not accustomed to, or could not use, the improved audibility of higher frequency consonantal information that was made available within the brief duration of this study (Thomas and Sparks, 1971).

Presentation of clipped speech at highest comfortable level in quiet also did not contribute to better perception of vowels when compared with unclipped speech at highest comfortable level. This lack of improvement in vowel perception between listening conditions was to be expected. The loudness of the rms of vowel peaks was likely the basis for highest comfortable level judgements so vowels were presented at near-equal levels between the two listening conditions. That is, if subjects set attenuator settings for highest comfortable level under the clipped and unclipped conditions based on the loudness of the vowel peaks, then it is to be expected that vowels would be perceived with the same accuracy under these two conditions – and that expectation agrees with the results obtained in this study.

Group mean recognition of initial consonants was significantly more accurate than that of final consonants. The better perception of initial consonants is likely due to clearer articulation of initial consonants as compared to final consonants. In addition, forward masking of the final consonant by the preceding high intensity vowel may reduce intelligibility.

The difference in scores in the recognition of initial as compared to final consonants was only 3.1 percentage points (Section 4.4). This finding is consistent with Pickett et al. (1972) who reported that subjects with moderately-severe to profound hearing loss had "somewhat better" recognition of initial than final consonants. Further review of their data reveals, however, that the group of subjects with severe to profound hearing loss recognized initial consonants at about the same level of accuracy as final consonants (less than one percentage point difference in scores).

Bode and Kasten (1971), studying normally hearing subjects, reported that an increase in clipping distortion was accompanied by greater drop in recognition of final consonants than of initial consonants. The difference in recognition scores, depending on level of distortion, was 10 to 19 percentage points. This difference is higher than that observed in the present study and may be explained as follows. Profound hearing loss is typically accompanied by much lower speech-perception scores than is normal hearing. These low scores may cause differences in the effects of clipping distortion between the perception of initial and final consonants to appear insignificant. Profound hearing loss is also typically accompanied by distortion caused by degraded spectral resolution and low-pass filtering resulting from a sloping loss. This distortion within the auditory system may mask the effects of distortion caused by clipping. The result may be that clipping distortion has less of an effect on scores for perception of initial and final consonants by subjects with profound hearing loss than by subjects with normal hearing.

Presentation level itself may also contribute to distortion within the ear. With high presentation levels, the cochlea may be pushed to function at levels at which it was not intended to operate. These internal nonlinearities, as described by Tonndorf (1986) and others, may influence the outcomes of research such as the present study.

5.3. Issues relating to internal validity

5.3.1. The repeatability of speech-awareness, highest-comfortable and sensation levels

The repeatability of speech-awareness thresholds and highest comfortable levels across test sessions was initially of concern. While the group mean settings for speech-awareness thresholds and highest comfortable level were very similar across replications, there was considerable variability within individual subjects. The standard deviation of 3.2 dB (Table 4.7) in the present study for speech-awareness thresholds measurements was, however, consistent with reports by other researchers as reviewed by Skinner (1988, p. 132-136).

These studies included subjects with hearing loss listening to pure tones or narrow-band noise. Witting and Hughson (1940) reported a standard deviation of 3.8 dB in thresholds measured many times over a period of two to 28 months. Skinner, Miller, DeFilippo, Dawson and Popelka (1986) reported a standard deviation of 2.7 dB for thresholds measured one day apart, Matsumoto (1983) reported 3.2 dB for thresholds measured seven days apart, and Skinner and Miller (1983) reported 2.7 dB for thresholds measured one day apart.

The standard deviation of 3.0 dB (Table 4.7) for highest comfortable level was, again, consistent with other studies that included subjects with hearing loss. These other studies, however, measured variability over time in other loudness judgements such as most comfortable levels or uncomfortable levels. The measurements of most comfortable and uncomfortable level, by definition, "bracket" the measurement of highest comfortable level as used in the present study and will be used here for comparison. For standard deviations in repeated measures of most comfortable levels, Cox and Bisset (1982) reported 3.9 dB measured over an unspecified length of time and with speech babble. Skinner and Miller (1983) reported a standard deviation of 4.1 dB measured over an unspecified number of days and with narrow-band noise. For standard deviations in uncomfortable-level measurements in response to narrow-band noise, Matsumoto (1983) reported 5.4 dB and Skinner and Miller (1983) reported 4.3 dB, both measured over an unspecified number of days. Thus, the variability observed in subjects' judgements of speech-awareness thresholds and highest comfortable level in the present study is in keeping with other reports in the literature.

The standard deviation of 4.2 dB (Table 4.7) for sensation level across replications was similar to results obtained by Pascoe, Miller, Skinner, Albee, Freiert and Hack (1980) as reviewed by Skinner (1988, p. 135). Pascoe et al. reported standard deviations in sensation level for hearing-impaired listeners responding to pure tones of between 3.2 and 7.2 dB as measured over an unspecified number of months. Thus, the variability in sensation level in the present study is consistent with a previous report.

The 95 percent confidence limits for speech-awareness threshold of +/- 6.8 dB and +/- 6.3 dB for highest comfortable level (Table 4.7) illustrate the necessity of permitting subjects to choose speech-awareness thresholds and highest comfortable level across sessions. If use of the same settings from the first session had been required in the second session and these thresholds had truly changed (not just the subjects' interpretation of the instructions for reporting these thresholds), softer speech sounds audible in the first session could have become inaudible in the second. The effect could have been to artificially lower scores in the second session. Conversely, testing could not have proceeded if levels judged as comfortable in the first session were judged as too loud in the second session.

5.3.2. Sensitivity

Because there was no significant difference between the intelligibility of clipped and unclipped speech when both are at highest comfortable level, it is necessary to consider the power of the design. A power analysis was used based on 95 percent confidence limits for differences between mean scores. The standard error of the difference between two scores is given by:

$$se = 100sqrt(p(1-p)2/(ns)) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where:

se = standard error of the difference in percentage points

p = mean of correct responses for clipped and unclipped speech combined, at highest comfortable level

n = number of items contributing to the score of a single subject

s = number of subjects

The group mean phoneme-recognition score under the two conditions at highest comfortable level was 17.8 percentage points (17.8% equals the average of 18.2% and 17.4% from Table 4.2). The effective number of test items was estimated in the following manner. Each of the three listening conditions presented four lists and each list contains thirty phonemes, for a total of 120 phonemes. However, according to Boothroyd and Nittrouer (1988), each list behaves as if it consists of twenty five independent phonemes, slightly less than the actual count of thirty. Therefore, it is estimated that there were approximately 100 independent items tested under each condition - four lists times 25 independent items per list. Thus, n equals 100. The number of subjects with profound hearing loss was 16 (i.e., $s = 16$).

Therefore:

$p = 0.178$

$n = 100$

$s = 16$

Substituting in equation (1) gives:

se = 1.35 percentage points

The value of "t" for 95 percent confidence limits is approximately 2.0. It may be concluded, therefore, that the design is sensitive to a difference of group means in the region of 2.7 percentage points (product of se of 1.35 and t of 2.0). This is equivalent to one phoneme in a list. It can be concluded, then, that differences in phoneme-recognition scores between conditions at highest comfortable level, if they exist, are smaller than 3.0 percentage points.

5.3.3. Order of listening conditions

One potential issue regarding validity was the consistent and repetitive order of presentation of the three listening conditions. This aspect of the design was necessary because of testing-time constraints and an effort to avoid some orders of listening conditions that would elicit negative reactions from subjects. Subjects' familiarity with the task could possibly have improved scores as they progressed through the conditions, yielding better scores under the later conditions. It should be noted, however, that the tasks were very similar to routine clinical testing familiar to the subjects. Any experience-based improvement should have already occurred. The extent to which a subject may correctly or incorrectly perceive words that are processed in a manner very similar to their everyday experience with hearing aids does not seem to readily lend itself to any new changes in adaptation, sensitization or learning that may affect performance across test conditions. Boredom could have caused progressively poorer scores under the later conditions. Judging from the reasonable level of subject interest in the experimental process, boredom did not appear to have been a problem. There are grounds, therefore, to believe that order effects did not occur but there is no way to be sure without further research.

Randomization of the order of the three listening conditions could have addressed concerns about order effects. There would still be the possibility, however, that experience with one condition would affect scores measured under another condition and the effect would not be reciprocal. Thus, randomization itself would not necessarily prevent order effects.

5.3.4. Acclimatization

There was the possibility that, had subjects had been allowed to become more accustomed to the listening conditions or even provided with training, the results would have been different. There are, however, changes in listening conditions experienced daily by these subjects. Changes in talker characteristics, room acoustics, and amplification equipment alter the acoustic information available to the hearing-aid user. In other words, these subjects were used to working with varying acoustic signals. It seems likely, but not proven, that acclimatization would not have a large effect on results.

5.3.5. Potential errors

Errors in equipment settings were minimized or eliminated by repetitive checks before, during and after test sessions. Written, detailed records of settings were kept at all times and changes in settings were limited to those described in the test procedures. Subjects' response sheets were scored once and then again by the same set of rules. These rules were designed and previously tested for consistent application. In the very few incidences where responses were scored twice and two different scores were obtained, a subject's written response was re-analyzed according to these rules and the correct score was determined. Only one scorer (the experimenter) was used to further maintain consistency and reduce the potential for errors. Data entry was checked and all calculations were repeated and sometimes in different ways in order to uncover potential errors.

5.4. Issues relating to external validity

The subjects participating in this study used primarily an auditory/oral approach to communication and, as such, were trained to use their hearing. The relevancy of their results is probably confined to profoundly deaf persons who are similarly trained. It is likely that results would be different for persons who use other modes of communication. These subjects were also within the age range of 14 to 16 years. It remains to be seen if results from this study apply to persons of much different ages.

The task of phoneme recognition is different from everyday comprehension of connected speech. The task is actually harder – little contextual information is provided, speech reading is not possible, and no help is available from sentence intonation. On the other hand, subjects need only to attend to one word at a time and the processing capacity for message comprehension is not required. It is reasonable to assume, however, that a monotonic, correlational relationship exists between phoneme recognition and performance at the sentence level. That is, as phoneme recognition improves, so does performance at the sentence level. In fact, this relationship has been demonstrated with normally-hearing (Boothroyd and Nittrouer, 1988) and hearing-impaired subjects (Boothroyd and Iglehart, 1997). It is, therefore, reasonable to generalize findings based on phoneme recognition to more realistic speech stimuli.

The negative affects of peak-clipping distortion on subjective judgements of quality of a speech signal are well known for normally hearing and moderately to severely hearing-impaired listeners but are not well documented for listeners with profound hearing loss. In the present study, subjects volunteered remarks that compared conditions #1 and #2 and compared conditions #2 and #3. A subjective comparison of quality between conditions #1 and #3 may have been advisable in order to determined if, for subjects with profound hearing loss, the benefits of peak clipping to speech perception are offset by an unacceptable loss of quality. On the other hand, these subjects may have had hearing that was too poor for them to sense a significant difference in quality between conditions #1 and #3.

Only one level of clipping was used in the present study. In theory, this was enough to significantly reduce the effective dynamic range of the speech signal without introducing excessive distortion. Distortion is considered here to be excessive when it outweighs the benefit of the additional gain that clipping allows to be introduced to the signal without exceeding listener comfort. This level of clipping was, however, chosen by informed guesswork. It is quite possible that the optimal amount of clipping (i.e., greatest benefit with the least cost) would be higher, or lower, than 15 dB. In everyday use of amplification, of course, it is the clipping level that is fixed, not the amount of clipping. In other words, the amount of clipping varies with the input level (no clipping with low inputs, high clipping levels with high inputs). These issues are in need of further research.

Hearing aids have narrower bandwidth and more frequency shaping than the broad-band and flat-response system used in the present study. To what extent the amplification system used in this experiment replicates the acoustic signal of personal amplification systems is open to question. Results of an earlier experiment (Boothroyd and Iglehart, 1997) indicated that there were no differences in speech-perception scores when subjects with very severe or profound hearing loss listened either with their personal aids or with the amplification system used for the present study. That study, however, did not include peak clipping. Previous research has shown that high-pass filtering in conjunction with clipping can influence speech perception – at least in normally hearing subjects in noise. Moreover, some of the supposed benefits of clipping may also be obtained through high-frequency emphasis in amplification. Perception with profound hearing loss, on the other hand, which is typically poorer in the high frequencies, may not benefit from improved access to high-frequency cues. It is possible, nevertheless, that the benefits of clipping may be less apparent with hearing-aid amplification – either simulated or real.

The acoustic conditions of listening through headphones were ideal for listening – quiet and no reverberation – and quite unlike real listening environments. It is known that noise combined with speech prior to clipping results in intermodulation distortion and an increase in the effective level of noise. Moreover, previous research has shown that the introduction of noise into processing of speech signals reduces the benefit of clipping. It may well be, therefore, that the present findings apply only in quiet. On the other hand, the benefit to perception of self-generated speech might still be expected to apply due to the high speech-to-noise ratio associated with such short talker-receiver distance. These issues are in need of further research.

It is useful to consider under what listening conditions the output of an aid in everyday use might exceed the limiting level adjusted to protect residual hearing. The average level of speech in quiet environments at conversational distance is typically 55 dB(Leq) (Pearsons et al., 1976, p.29). The typical maximum gain available from a powerful hearing aid before the onset of acoustic oscillation is about 60 dB. As a result, conversations in quiet may typically be amplified to an output of approximately 115 dB(Leq). In order to compare this data to the present study, it is necessary to determine the Leq levels of speech stimuli used in the present study. These levels are provided in Appendix G for clipped and unclipped speech and are approximately 4 dB below the average vowel peak levels (123 dBSPL) or approximately 119 dB(Leq). These data indicate that, while amplified conversational speech in quiet may not typically exceed these subjects' highest comfortable levels, average vowel peaks of louder speech will likely exceed the limiting adjustments in powerful hearing aids. These louder conversational sounds may be as common as speech raised above noisy environments (an average of 67 dB(Leq); Pearsons et al., p. 29) or the hearing-aid user's own voice.

5.5. Further research

The findings of the present study raise several questions important to the fitting of amplification for subjects with profound hearing loss. What is the optimal level of clipping? Higher levels of gain will increase the power of the signal but will also undoubtedly cause more distortion. The clipping level that offers the best balance between benefits of additional power and costs of distortion have not been established for subjects with profound hearing loss.

What are the effects of noise that is combined with a speech signal prior to clipping and at different speech-to-noise ratios? Noise that is clipped with the speech signal, as discussed earlier, leads to intermodulation distortion. The specific effects of interactions of noise at different speech-to-noise ratios and clipping on speech perception by these subjects are not well known.

What are the effects of frequency filtering and how does it interact with the factors of clipping and noise? Effects of the filtering parameters of frequency cutoff and slope most beneficial to speech perception by profoundly deaf subjects are not well established.

Will acclimation to a listening condition significantly alter perception of speech within that condition? The benefit of any speech processing design may depend on the extent of listening experience with that design. The role of acclimatization in the evaluation of hearing-aid design for the profoundly hearing impaired is not well understood.

To what extent do the benefits of clipping carry over to perception by speech reading combined with audition? There is the possibility that the benefits of clipping will be proportional to speech-perception scores and, therefore, insignificant to the low auditory-alone perception scores that are typical of subjects with profound losses. The benefit of clipping may, therefore, be underestimated if evaluated by audition alone. In comparison, speech perception by audition and speech reading combined typically yields higher scores than auditory-alone perception. Clipping may show significant proportional improvement in scores obtained with audition and speech reading combined. Testing by audition and speech reading combined, therefore, may give a more realistic and higher estimate of the benefits of clipping for this population.

The amplitudes in Appendix G, the subjects' speech-awareness levels and highest comfortable levels all reflect the lower frequencies of the speech signal – primarily vowel peaks in the frequency range of 500 to 1000 Hz. Because of sloping audiograms, high-frequency thresholds are likely to be much poorer and the sensation levels in the high frequencies much lower than is indicated in the present study. It is quite possible that much of the fricative and burst information and perhaps higher vowel formants (F3 and F4) were inaudible under conditions #2 and #3 for many subjects. Further studies could take into account the frequency dependence of long-term amplitudes of speech signals in relation to pure tone thresholds.

There was evidence in the present study that loudness judgements by subjects with profound hearing loss may not depend only on level. A large portion of the subjects chose loudest comfortable levels for backward-playing speech that were lower than for forward-playing speech. Subjects also described the backward-playing speech as sounding unusual. Further research may indicate whether subjects with profound hearing loss may more cautiously judge the loudness of speech that they do not trust.

5.6. Clinical implications

The findings of the present study have several clinical implications. If an existing hearing aid has a saturation sound pressure level that is too low to provide optimal intelligibility without clipping for the user with profound hearing loss, there are two options. One option is to use an aid with greater output. The other is to increase gain beyond the point at which clipping begins. The results of the present study indicate that the two options will provide equal benefit – at least in quiet. A small advantage of a less powerful aid is that it may cost slightly less. Even if the choice is made to adjust a lower-output aid so that conversational speech is not clipped, the user's voice may drive the system into clipping. The results of the present study suggest that this is not necessarily a problem for the user.

A possible advantage of an aid with higher saturation sound pressure level is the lack of distortion. Distortion, however, as reported in this and previous studies may not necessarily be problematic for a listener with profound hearing loss. At the same time, higher saturation sound pressure level has an important drawback. The user's hearing is not protected from excessive loudness since the saturation sound pressure level is set some 10 dB above the user's highest comfortable level. Many environmental sounds and self-generated speech will be uncomfortable and may damage the listener's residual hearing. The hearing-aid user, especially if a child, will likely turn down the volume of the aid and miss otherwise useful speech cues.

5.7. Conclusions

For these subjects with profound hearing loss who used auditory/oral communication:

1. Clipped speech was more intelligible than unclipped speech in quiet when both were presented at the same instantaneous peak level (18.2 percent versus 13.1 percent, respectively).
2. Unclipped speech was not more intelligible than clipped speech when both were presented in quiet and at highest comfortable level (17.1 percent versus 18.2 percent, respectively).

These results indicate that, in fitting of hearing aids if highest comfortable level has not been attained, it may be appropriate for persons with profound hearing loss to raise the gain past the point at which clipping begins in order to raise the rms level of the speech signal, thereby increasing intelligibility. The resulting distortion, at least in quiet, may not decrease intelligibility.

Appendix A. The following two pages are the consent forms signed by parents and subjects. The dates included at the top of each page were the approval dates provided by the Institutional Review Board of the City University of New York.

Approved: April 20, 1998
Expires : March 21, 1999

CLARKE

The Clarence W. Barron Research Department

Clarke School for the Deaf
Center for Oral Education

Founded 1867

Yvonne Mullen, Ed.D., NCSP
Director, Psychology Division

Effie G. Ronald, C.A.G.S.
School Psychologist

Frank Iglehart, M.A., CCC-A
Acting Director, Hudgins Audiology Division

Don Salvucci, M.Ed., Ed.M., CFY-A
Educational Audiologist

Date:

Parents' names and address

Dear Mr. and Mrs. ,

We are planning a study of hearing aids and classroom FM systems at the Clarke School and would like your permission for (child's name) to participate. The study will be conducted in a classroom setting and in the Audiology Department's soundbooth and we shall compare different types of behind-the-ear hearing aid and classroom FM amplification. The task of the students will be to listen to, and identify, words and syllables. The overall purpose will be to examine the relative benefits to the speech perception of deaf students provided by various types of linear and compression amplification.

There are no known risks involved. The aids and FM receivers will be adjusted according to the needs of each participant and all signals will be at comfortable levels. The study will involve a total of about 3 hours of testing, spread over about 6 sessions during the course of the school year. It will not interfere with class time. The results of the study may be published and reported at scientific meetings but individual participants will not be identified. Please note that participation is entirely voluntary on the part of both parents and students, and any permission given now may be withdrawn at any time without penalty. When the study is completed, we plan to share the results with all Clarke parents.

This study is funded by the National Institute of Disability Rehabilitation and Research and sponsored by the City University of New York (CUNY). If you have any questions or would like more information, please contact Frank at (413) 584-3450 or the CUNY Office of Research and University Programs Sponsored Research at (212) 642-2059.

Sincerely,

Frank Iglehart, M.A., CCC-A   Pamela Paskowitz, PhD, CCC-SLP

(Please Tear Off and Return to the Audiology Department)

To: Frank Iglehart,

We/I give our/my permission for our/my son/daughter, (child's name), to participate in a study of hearing aids and FM systems as described in your letter dated (date). It is understood that this permission may be withdrawn at any time without penalty.

Signature ______________________ Name ________________________ Date __________

Signature ______________________ Name ________________________ Date __________

Round Hill Road
Northampton, MA 01060-2199
413/584-3450 (V.TTY)
FAX 413/586-6644

Approved: April 20, 1998
Expires : March 21, 1999

To: (student's name)

Date:

Dear

We would like your help in a study of hearing aids and classroom FM systems. If you agree, you will be asked to write down words and syllables that you hear through different hearing aids and FM receivers. The hearing aids and FM receivers will be set to match your hearing and none of the sounds will be too loud. Your scores will not be shown to anyone else unless you agree. Even if you say "yes" now, you can change your mind later and say "no".

If you want to know more about this study, please feel free to ask me.

We are also writing to your parents for permission. You can help us only if they also agree.

Sincerely,

Frank Iglehart, M.A., CCC-A

(Please Tear Off and Return to the Audiology Department)

I agree to help with the study of hearing aids and FM systems. I know that I can stop participating at any time during the study.

Signature ______________________ Name ______________________ Date ____________

Appendix B. Word lists (Boothroyd, 1984) randomly chosen for each listening condition and test session.

List 1
1. SHIP
2. RUG
3. FAN
4. CHEEK
5. HAZE
6. DICE
7. BOTH
8. WELL
9. JOT
10. MOVE

List 2
1. FISH
2. DUCK
3. PATH
4. CHEESE
5. RACE
6. HIVE
7. BONE
8. WEDGE
9. LOG
10. TOMB

List 3
1. THUG
2. WITCH
3. TEAK
4. WRAP
5. VICE
6. JAIL
7. HEN
8. SHOWS
9. FOOD
10. BOMB

List 4
1. FUN
2. WILL
3. VAT
4. SHAPE
5. WREATH
6. HIDE
7. GUESS
8. COMB
9. CHOOSE
10. JOB

List 5
1. FIB
2. THATCH
3. SUM
4. HEEL
5. WIDE
6. RAKE
7. GOES
8. SHOP
9. VET
10. JUNE

List 6
1. FILL
2. CATCH
3. THUMB
4. HEAP
5. WISE
6. RAVE
7. GOT
8. SHOWN
9. BED
10. JUICE

List 7
1. BADGE
2. HUTCH
3. KILL
4. THIGHS
5. WAVE
6. REAP
7. FOAM
8. GOOSE
9. NOT
10. SHED

List 8
1. BATH
2. HUM
3. DIG
4. FIVE
5. WAYS
6. REACH
7. JOKE
8. NOOSE
9. POT
10. SHELL

List 9
1. HUSH
2. GAS
3. THIN
4. FAKE
5. CHIME
6. WEAVE
7. JET
8. ROB
9. DOPE
10. LOSE

List 10
1. JUG
2. LATCH
3. WICK
4. FAITH
5. SIGN
6. BEEP
7. HEM
8. ROD
9. VOTE
10. SHOES
11.

List 11
1. MATH
2. HIP
3. GUN
4. RIDE
5. SIEGE
6. VEIL
7. CHOSE
8. SHOOT
9. WEB
10. COUGH

List 12
1. HAVE
2. WIG
3. BUFF
4. MICE
5. TEETH
6. JAYS
7. POACH
8. RULE
9. DEN
10. SHOCK

List 13
1. KISS
2. BUZZ
3. HASH
4. THIEVE
5. GATE
6. WIFE
7. POLE
8. WRETCH
9. DODGE
10. MOON

List 14
1. WISH
2. DUTCH
3. JAM
4. HEATH
5. LAZE
6. BIKE
7. ROVE
8. PET
9. FOG
10. SOON

List 15
1. HUG
2. DISH
3. BAN
4. RAGE
5. CHIEF
6. PIES
7. WET
8. COVE
9. LOOSE
10. MOTH

List 16
1. WAGE
2. RAG
3. BEACH
4. CHEF
5. DIME
6. THICK
7. LOVE
8. ZONE
9. HOP
10. SUIT

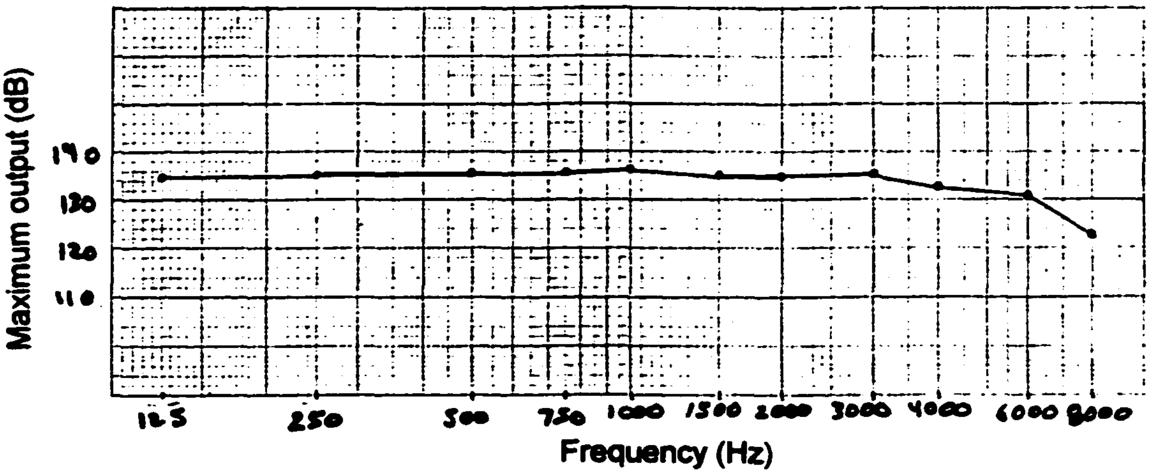
List 17
1. JADE
2. CASH
3. THIEF
4. SET
5. WINE
6. GIVE
7. RUB
8. HOLE
9. CHOP
10. ZOOM

List 18
1. SHAVE
2. JAZZ
3. THEME
4. FETCH
5. HEIGHT
6. WIN
7. SUCK
8. ROBE
9. DOG
10. POOL

List 19
1. VASE
2. CAB
3. TEACH
4. DEATH
5. NICE
6. FIG
7. RUSH
8. HOPE
9. LODGE
10. WOMB

List 20
1. CAVE
2. RASH
3. TEASE
4. JELL
5. GUIDE
6. PIN
7. FUSS
8. HOME
9. WATCH
10. BOOTH

Appendix C. **Frequency-response curve of a Stimulus Presentation System. Maximum undistorted acoustic output through TDH 50 headphones into a 6cc coupler as measured with the output attenuator adjusted to 136 dBSPL.**


Maximum output (dB)
140
130
120
110
125
250
500
750
1000
1500
2000
3000
4000
6000
8000
Frequency (Hz)

Appendix D. Written instructions provided to subjects at the beginning of each test session.

Your participation here is very important. You are helping to find ways to build better hearing aids and FM systems for people with severe and profound hearing losses.

When we start, you will hear speech sounds.

Please adjust the volume to the level that is the loudest comfortable level for you. The black knobs on the front of the silver-colored box are the volume controls.

Next, I will tell when:

> You will hear the sentence "Write the word _______." At the end of the sentences will be a different word each time. Please write down only the last word.
>
> If you are not sure of what you heard, that is OK. Please write down what you think you heard. If you have no idea what the word was, you may leave the space blank if you want to.
>
> Some of the sentences may later sound too soft. Please do not change your volume settings until I let you know that it is OK to change the volume.

Please feel free to ask me any questions. Please remember – there are no grades given and no teachers will see what you write.

Thank you for your help.

Appendix E. Scores by individual subjects for recognition of initial consonants, medial vowels, final consonants, and words in the three listening conditions. Table E1 provides scores for session 1. Table E2 provides scores for session 2. All scores are provided as percentage correct. Means, standard deviations and standard errors are calculated only for 16 subjects with profound hearing loss.

Table E1. Scores in session 1 by individual subjects for initial consonants (C1), medial vowels (V), final consonants (C2) and words. All scores are provided as percentage correct.

Subject	Hearing loss 3-freq. ave. (dB)	Listening condition (percent correct)											
		# 1				# 2				# 3			
		C1	V	C2	Words	C1	V	C2	Words	C1	V	C2	Words
s1	48	85.0	85.0	97.5	75.0	97.5	97.5	95.0	92.5	97.5	92.5	95.0	87.5
s2	60	77.5	85.0	82.5	67.5	80.0	87.5	75.0	55.0	87.5	90.0	87.5	75.0
s3	93	7.5	12.5	2.5	0.0	10.0	7.5	7.5	0.0	10.0	30.0	2.5	0.0
s4	95	55.0	50.0	25.0	15.0	15.0	35.0	12.5	0.0	35.0	35.0	10.0	5.0
s5	98	15.0	25.0	5.0	0.0	15.0	20.0	12.5	0.0	17.5	32.5	20.0	2.5
s6	98	2.5	10.0	0.0	0.0	2.5	12.5	0.0	0.0	5.0	10.0	0.0	0.0
s7	98	10.0	37.5	5.0	2.5	22.5	30.0	5.0	2.5	12.5	32.5	35.0	2.5
s8	98	12.5	25.0	5.0	2.5	7.5	12.5	5.0	0.0	10.0	12.5	10.0	0.0
s9	100	32.5	37.5	22.5	5.0	30.0	42.5	27.5	12.5	40.0	62.5	42.5	27.5
s10	100	12.5	12.5	7.5	0.0	7.5	5.0	5.0	0.0	5.0	15.0	7.5	0.0
s11	102	22.5	35.0	17.5	5.0	22.5	30.0	5.0	0.0	20.0	25.0	17.5	5.0
s12	102	12.5	17.5	10.0	0.0	2.5	10.0	0.0	0.0	2.5	15.0	0.0	0.0
s13	102	25.0	30.0	25.0	7.5	12.5	40.0	22.5	7.5	22.5	32.5	20.0	5.0
s14	105	5.0	30.0	10.0	0.0	5.0	17.5	2.5	0.0	7.5	22.5	5.0	0.0
s15	105	7.5	35.0	2.5	2.5	5.0	12.5	2.5	0.0	10.0	37.5	2.5	2.5
s16	105	7.5	5.0	2.5	0.0	2.5	2.5	5.0	0.0	5.0	12.5	7.5	0.0
s17	105	12.5	15.0	2.5	0.0	7.5	7.5	2.5	0.0	7.5	20.0	5.0	0.0
s18	108	7.5	20.0	7.5	2.5	7.5	10.0	7.5	0.0	7.5	15.0	10.0	0.0
Means for profound		15.5	24.8	9.4	2.7	10.9	18.4	7.7	1.4	13.6	25.6	12.2	3.1
sd		13.1	12.4	8.4	4.0	8.2	12.9	7.7	3.5	10.9	13.5	12.2	6.8
se		3.3	3.1	2.1	1.0	2.0	3.2	1.9	0.9	2.7	3.4	3.1	1.7

Table E2. Scores in session 2 by individual subjects for initial consonants (C1), medial vowels (V), final consonants (C2) and words. All scores are provided as percentage correct.

Subject	Hearing loss 3-freq. ave. (dB)	Listening condition (percent correct)											
		# 1				# 2				# 3			
		C1	V	C2	Words	C1	V	C2	Words	C1	V	C2	Words
s1	48	95.0	97.5	95.0	92.5	95.0	92.5	97.5	87.5	92.5	92.5	95.0	85.0
s2	60	90.0	92.5	77.5	70.0	80.0	95.0	80.0	67.5	85.0	95.0	82.5	72.5
s3	93	27.5	42.5	22.5	0.0	17.5	35.0	20.0	2.5	10.0	20.0	10.0	0.0
s4	95	35.0	47.5	27.5	15.0	12.5	22.5	5.0	0.0	17.5	47.5	12.5	10.0
s5	98	20.0	22.5	10.0	2.5	5.0	20.0	5.0	0.0	20.0	15.0	7.5	0.0
s6	98	2.5	2.5	2.5	0.0	12.5	15.0	10.0	0.0	12.5	12.5	12.5	0.0
s7	98	35.0	52.5	25.0	17.5	22.5	30.0	25.0	2.5	27.5	32.5	22.5	7.5
s8	98	5.0	20.0	15.0	2.5	10.0	22.5	5.0	2.5	7.5	12.5	7.5	0.0
s9	100	40.0	37.5	35.0	7.5	10.0	20.0	25.0	0.0	27.5	60.0	27.5	5.0
s10	100	20.0	35.0	17.5	2.5	5.0	22.5	10.0	0.0	15.0	25.0	15.0	2.5
s11	102	20.0	32.5	10.0	2.5	15.0	25.0	12.5	2.5	25.0	35.0	12.5	5.0
s12	102	5.0	12.5	5.0	0.0	12.5	12.5	5.0	2.5	15.0	22.5	5.0	5.0
s13	102	27.5	30.0	22.5	7.5	17.5	32.5	12.5	5.0	30.0	22.5	22.5	0.0
s14	105	2.5	25.0	10.0	0.0	10.0	25.0	7.5	2.5	10.0	12.5	5.0	0.0
s15	105	20.0	27.5	10.0	0.0	2.5	17.5	2.5	0.0	17.5	32.5	17.5	2.5
s16	105	7.5	15.0	5.0	0.0	2.5	10.0	7.5	0.0	2.5	20.0	7.5	0.0
s17	105	12.5	20.0	7.5	0.0	5.0	12.5	0.0	0.0	10.0	15.0	5.0	0.0
s18	108	7.5	10.0	7.5	0.0	10.0	10.0	7.5	0.0	5.0	10.0	10.0	0.0
Means for profound		18.0	27.0	14.5	3.6	10.6	20.8	10.0	1.3	15.8	24.7	12.5	2.3
sd		12.5	13.9	9.5	5.6	5.7	7.7	7.5	1.6	8.4	13.9	6.9	3.2
se		3.1	3.5	2.4	1.4	1.4	1.9	1.9	0.4	2.1	3.5	1.7	0.8

Appendix F. Listening levels (attenuator settings in dB) for Speech-Awareness Thresholds, Highest Comfortable Levels and Sensation Levels. Data are provided for each session and averages of the two sessions. Table F1 provides data for listening condition #1. Table F2 provides data for listening condition #3. Means, standard deviations and standard errors are calculated only for the 16 subjects with profound hearing loss.

Table F1. Measures for listening condition #1 of Speech-Awareness Thresholds (SAT), Highest Comfortable Levels (HCL) and Sensation Levels (SL) for each subject. Data are provided for each session and averages of the two sessions.

Subj	Hearing loss 3-freq. ave. (dB)	Listening condition # 1								
		Session 1			Session 2			Average of 1 and 2		
		SAT dBSPL	HCL dBSPL	SL dB	SAT dBSPL	HCL dBSPL	SL dB	SAT dB	HCL dB	SL dB
s1	48	52	92	40	50	102	52	51	97	46
s2	60	66	106	40	74	114	40	70	110	40
s3	93	94	122	28	96	128	32	95	125	30
s4	95	100	120	20	100	120	20	100	120	20
s5	98	84	122	38	86	116	30	85	119	34
s6	98	96	118	22	94	128	34	95	123	28
s7	98	90	122	32	88	122	34	89	122	33
s8	98	98	124	26	96	122	26	97	123	26
s9	100	92	134	42	98	122	24	95	128	33
s10	100	94	112	18	92	128	36	93	120	27
s11	102	96	124	28	94	122	28	95	123	28
s12	102	106	124	18	98	130	32	102	127	25
s13	102	106	134	28	106	132	26	106	133	27
s14	105	114	136	22	108	130	22	111	133	22
s15	105	106	126	20	110	130	20	108	128	20
s16	105	108	120	12	106	124	18	107	122	15
s17	105	98	122	24	100	126	26	99	124	25
s18	108	114	132	18	114	134	20	114	133	19
Means for profound		99.8	124.5	24.8	99.1	125.9	26.8	99.4	125.2	25.8
sd		8.3	6.3	7.6	7.6	4.8	5.6	8.0	4.7	5.5
se		2.1	1.6	1.9	1.9	1.2	1.4	2.0	1.2	1.4

Table F2. Measures for listening condition #3 of Speech-Awareness Thresholds (SAT), Highest Comfortable Levels (HCL) and Sensation Levels (SL) for each subject. Data are provided for each session and averages of the two sessions.

Subj	Hearing loss 3-freq. ave. (dB)	Listening condition # 3								
		Session 1			Session 2			Average of 1 and 2		
		SAT dBSPL	HCL dBSPL	SL dB	SAT dBSPL	HCL dBSPL	SL dB	SAT dB	HCL dB	SL dB
s1	48	62	94	32	64	108	44	63	101	38
s2	60	76	122	46	82	122	40	79	122	43
s3	93	114	130	16	102	132	30	108	131	23
s4	95	106	126	20	110	128	18	108	127	19
s5	98	104	134	30	90	130	40	97	132	35
s6	98	108	136	28	102	134	32	105	135	30
s7	98	104	124	20	98	130	32	101	127	26
s8	98	110	134	24	106	126	20	108	130	22
s9	100	106	136	30	110	136	26	108	136	28
s10	100	106	124	18	104	132	28	105	128	23
s11	102	102	130	28	108	132	24	105	131	26
s12	102	112	130	18	112	134	22	112	132	20
s13	102	114	136	22	114	136	22	114	136	22
s14	105	126	136	10	118	136	18	122	136	14
s15	105	114	132	18	116	134	18	115	133	18
s16	105	118	134	16	110	134	24	114	134	20
s17	105	106	130	24	110	132	22	108	131	23
s18	108	126	136	10	126	136	10	126	136	10
Means for profound		111.0	131.8	20.8	108.5	132.6	24.1	109.8	132.2	22.4
sd		7.4	4.3	6.3	8.4	3.0	7.2	7.3	3.2	6.0
se		1.8	1.1	1.6	2.1	0.7	1.8	1.8	0.8	1.5

Appendix G. Distributions of rms amplitude of the speech stimuli in the three listening conditions are provided. To aid in the interpretation of the behavioral results, key parameters of the distribution of amplitudes in the speech stimuli were derived. The analyses were performed on the first three word lists (i.e., 30 words; Appendix B) for the two stimulus types (clipped and unclipped). For each of the 30 words the rms amplitude was first computed as a function of time. The calculation was done for successive 100 msec. samples using a 50 percent overlap. The peak value of rms amplitude was identified for each word. From the 30 peak values, a maximum, mean and a minimum were found. These measures clearly applied to the peak rms levels of the vowels within the words.

Three additional parameters were derived for the complete set of 30 words. To derive these parameters, the 30 words were concatenated – after removal of interword silences. An overall rms amplitude was then calculated. This measure provided the Leq rms level of the test material. Using, again, a 100 msec window, the rms amplitude was calculated as a function of time. A distribution of the resulting amplitude measures was prepared from which the modal rms level was determined. Additionally, the rms level below which 20 percent of the amplitude measures fell was determined. This criterion was chosen empirically to be 30 dB below the maximum peak vowel level for the unclipped speech. The choice was based on the known 30 dB dynamic range of speech and was an attempt to define the lowest intensity at which useful information might be found. Finally, the peak instantaneous level was converted to decibels. This last measure represents the rms level of a square wave whose peak amplitude is equal to that of the speech material (equivalent to saturation sound pressure level in a peak-clipping hearing aid).

The seven parameters, in descending order of magnitude, were:

1. Instantaneous maximum
2. Maximum of the peak vowel levels within words.
3. Average of the peak vowel levels within words.
4. Minimum of the peak vowel levels within words.
5. Modal value of rms amplitude of the 30 words.
6. Long-term equivalent rms amplitude of the 30 words.
7. Rms amplitude at the 20th percentile of the amplitude distribution.

Figure G1 shows the amplitude distributions of the clipped and unclipped speech material. The amplitude scale has been adjusted to represent the output under headphones when using the Stimulus Presentation System of the present study with the attenuator set at maximum (136 dBSPL). The values of the seven key parameters are as marked.

In Figure G2, the amplitudes have been readjusted to represent the distributions as they would occur under headphones with the attenuator set at the mean values shown in Table 4.6. The threshold values are estimates based on sensation level data of Table 4.6 and the assumption that judgements of speech awareness threshold were based on the average of the peak vowel levels.

These data are consistent with the conclusion that the subjects, on average, had full access to the useful information in the speech signal when listening at highest comfortable level to the clipped and unclipped speech (conditions #1 and #3). They are also consistent with the conclusion that a significant portion of the useful information fell below threshold under (unclipped) condition #2.

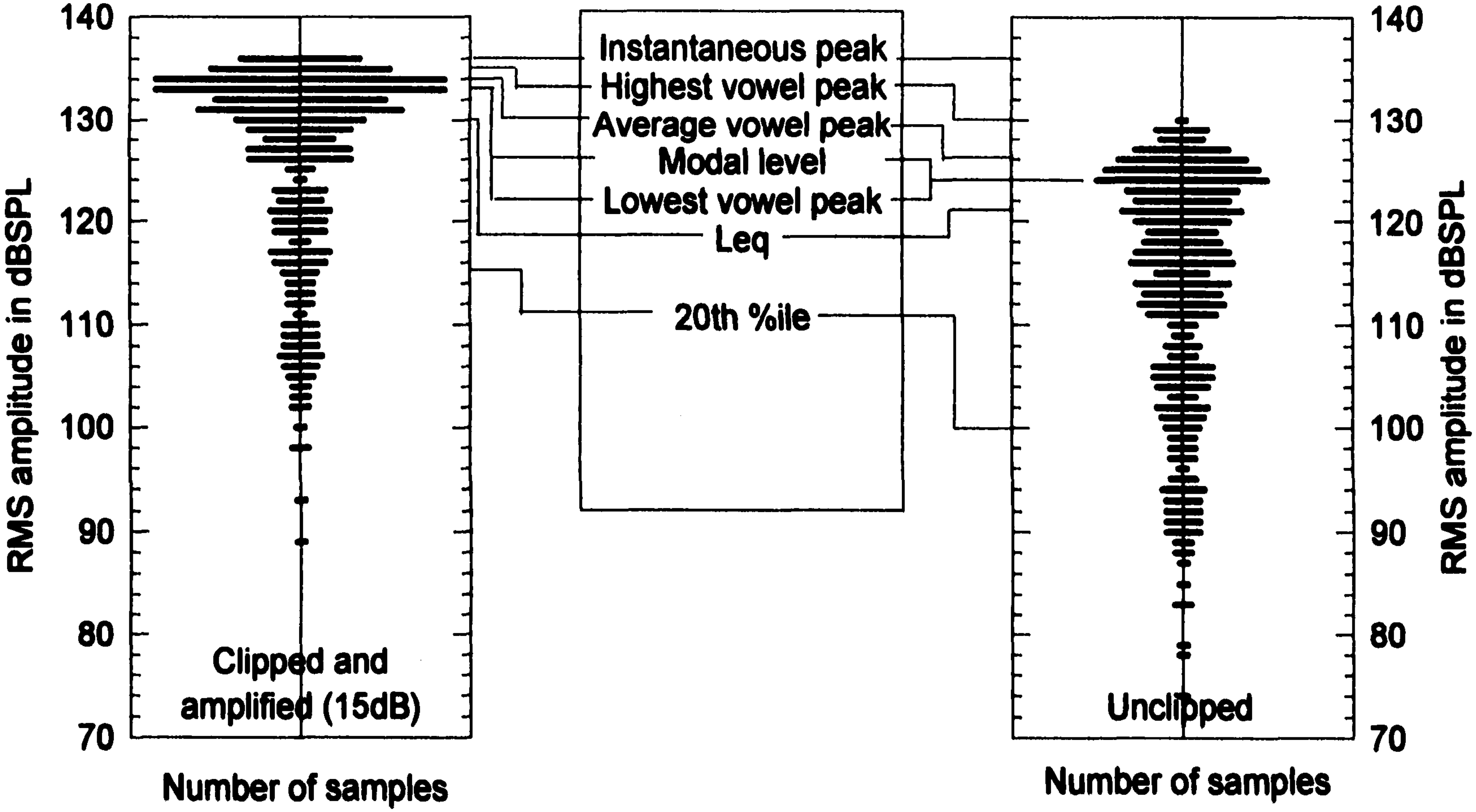


Figure G.1. Distribution of rms amplitude (integrated over 100 msec) in the clipped and unclipped speech stimuli. Amplitude is expressed in term of maximum output via TDH 50 earphones into a 6 cc coupler.

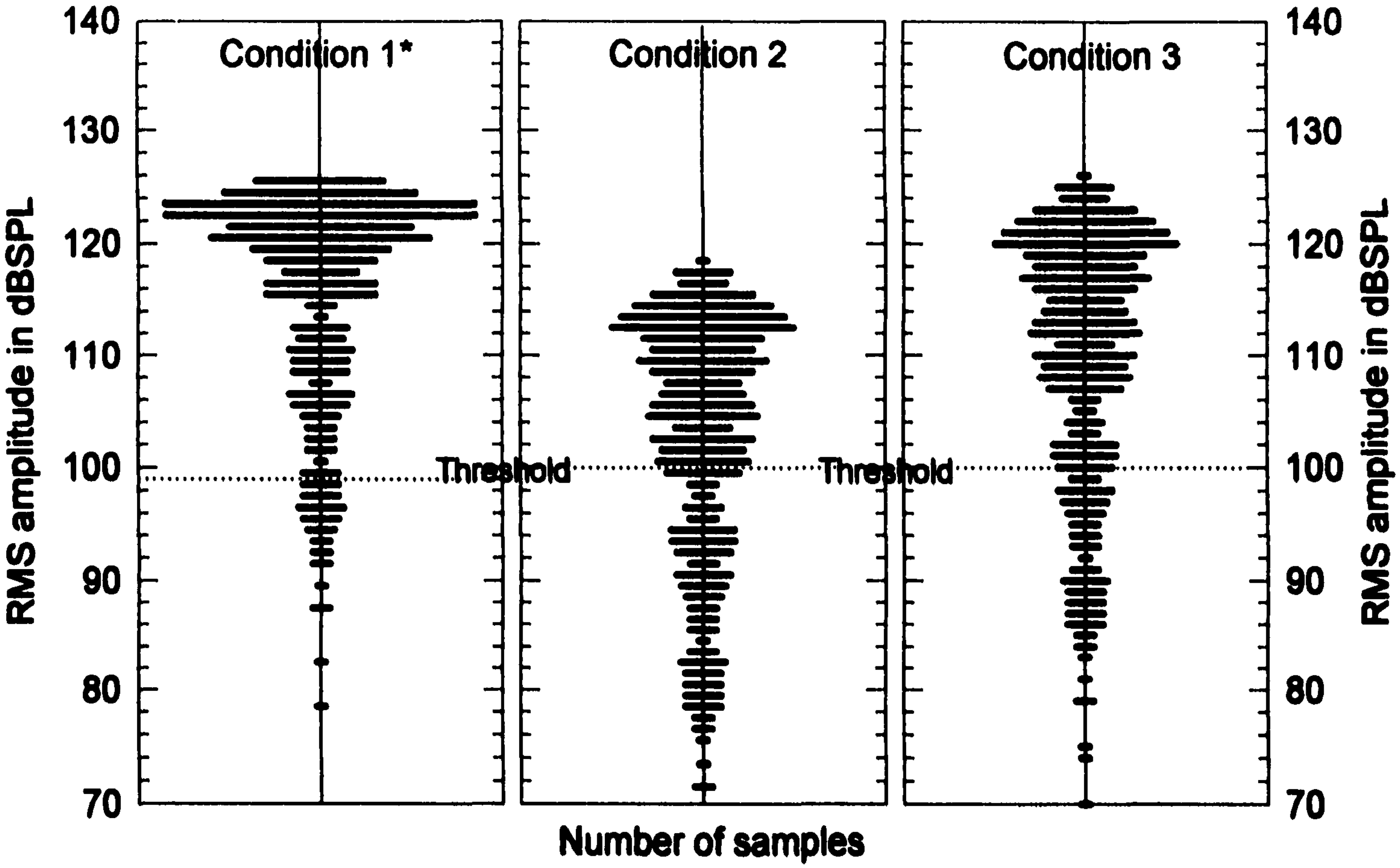


Figure G.2. Distribution of rms amplitude (integrated over 100 msec) in the clipped and unclipped speech stimuli. Amplitude levels are adjusted to match the group mean settings for the three listening conditions. Thresholds were estimated by taking the average rms vowel peak at HCL and subtracting the group mean sensation levels for the clipped and unclipped conditions.

*Condition 1 = speech clipped by 15dB and presented at Highest Comfortable Level (HCL)
Condition 2 = unclipped speech presented with the same instantaneous peak level as in Condition 1
Condition 3 = unclipped speech presented at HCL.

BIBLIOGRAPHY

Ainsworth WA (1967) Relative intelligibility of different transforms of clipped speech. *J Acous Soc Am*, 41 (5): 1272-1276.

ANSI (1960) Methods for Measurement of Electroacoustic Characteristics of Hearing Aids. *ANSI S3.3-1960*. New York: American National Standards Institute.

ANSI (1969) American National Standard Methods for the Calculation of the Articulation Index. *ANSI S3.5-1969*. New York: American National Standards Institute.

ANSI (1989) Specifications for Audiometers. *ANSI S3.2-1989*. New York: American National Standards Institute.

ANSI (1996) Specifications for Audiometers. *ANSI S3.6-1996*. New York: American National Standards Institute.

ANSI (1997) Methods for Measuring the Intelligibility of Speech Over Communication Systems. *ANSI S3.5-1997*. New York: American National Standards Institute.

Bench J, Kowal A and Bamford J (1979) The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Brit J Audiol*, 13: 108-112.

Bode DL and Kasten RN (1971) Hearing aid distortion and consonant identification. *J Acous Soc Am*, 14 (2): 323-331.

Boothroyd A (1968) Selection of hearing aids in children. Unpublished thesis, University of Manchester, England.

Boothroyd A (1984) Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *J Sp Hear Res*, 27: 134-144.

Boothroyd, A. (1993a) Speech perception, sensorineural hearing loss, and hearing aids. In Studebaker GA and Hochberg I (Eds.) *Acoustic Factors Affecting Hearing Aid Performance, 2nd Ed.*, Boston, Allyn and Bacon: 277-299.

Boothroyd A (1993b) Speech perception tests and hearing-impaired children. In Plant G and Spens KE (Eds.) *Profound Deafness and Speech Communication*, San Diego, CA, Singular Publishing Group Inc.: 345-371.

Boothroyd A, Erickson FN and Medwetsky L (1994) The hearing aid input: A phonemic approach to assessing the spectral distribution of speech. *Ear Hear*, 15 (6): 432-442.

Boothroyd A and Iglehart F (1997) Speech perception and method of amplification. Unpublished. Northampton, MA, Clarke School for the Deaf.

Boothroyd A and Iglehart F (1998) Experiments with classroom FM amplification. *Ear Hear*, 19(3):202-217.

Boothroyd A and Nittrouer S (1988) Mathematical treatment of context effects in phoneme and word recognition. *J Acous Soc Am*, 84: 101-114.

Cox RM and Bisset JD (1982) Prediction of aided preferred listening levels for hearing aid gain prescription. *Ear Hear*, 3: 66-71.

Cox RM and McDaniel DM (1984) Intelligibility ratings of continuous discourse: Application to hearing aid selection. *J Acous Soc Am*, 76 (3): 758-766.

Crain TR and Van Tasell DJ (1994) Effect of peak clipping on speech recognition threshold. *Ear Hear*, 15 (6): 443-453.

Dadson RS and King JH (1978) A determination of the normal threshold of hearing and its relation to the standardization of audiometers. *J Laryng Otol*, 46: 366-378.

Dawson P, Dillon H and Battaglia J (1990) Output limiting compression for the severe-profoundly deaf. *Austral J Audiol*, 13 (1): 1-12.

Dillon H and Storey L (1998) The National Acoustic Laboratories' procedure for selecting saturation sound pressure level of hearing aids: Theoretical Derivation. *Ear Hear*, 19 (4): 255-266.

Dreschler WA (1988) Dynamic-range reduction by peak clipping or compression and its effects on phoneme perception in hearing-impaired listeners. *Scand Audiol*, 17:45-51.

Dunn HK and White SD (1940) Statistical measurements on conversational speech. *J Acous Soc Am*, 11: 278-288.

Egan JP (1944) *Articulation Testing Methods II, Report OSRD No. 3802*, PB 22848, Psycho-Acoustic Laboratories, Publications Board, U.S. Department of Commerce, Washington D.C.

Egan JP (1948) Articulation testing methods. *Laryngoscope*, 58: 955-991.

Erber N and Alencewicz CM (1976) Audiologic evaluation of deaf children. *J Sp Hear Dis*, 41: 256-267.

Fletcher H (1950) *Speech and Hearing*. New York, D. Van Nostrand Co. Inc; 273-276.

Flynn MC, Dowell RC and Clark GM (1998) Aided speech recognition abilities of adults with a severe or severe-to-profound hearing loss. *J Sp Hear Res*, 41: 285-299.

Fortune TW and Preves DA (1992). Hearing aid saturation and aided loudness discomfort. *J Sp Hear Res*, 35: 175-185.

French NR and Steinberg JC (1947) Factors governing the intelligibility of speech. *J Acous Soc Am*, 19: 90-119.

Gionnini LA and Franzen RL (1978) Comparison of the effects of hearing aid distortion on performance scores for the MRHT and PB-50 tests. *J Aud Res*, 18: 203-208.

House AS, Williams CE, Hecker MHL, and Kryter KD (1965) Articulation testing methods: Consonantal differentiation with a closed-response set. *J Acous Soc Am*, 37: 158-166.

Hudgins CV, Marquis RJ, Nichols RH, Peterson GE, and Ross DA (1948) The comparative performance of an experimental hearing aid and two commercial instruments. *J Acous Soc Am*, 20 (3): 241-258.

Kates JM and Kozma-Spytek L (1994) Quality ratings for frequency-shaped ands peak-clipped speech. *J Acous Soc Am,* 95 (6): 3586-3594.

Kozma-Spytek L, Kates JM and Revoile SG (1996) Quality ratings for frequency-shaped peak-clipped speech: Results for listeners with hearing loss. *J Sp Hear Res*, 39: 1115-1123.

Kreul RE, Nixon JC, Kryter KD, Bell DW, Lang JS and Schubert ED (1968) A proposed clinical test of speech discrimination. *J Sp Hear Res*, 11: 536-552.

Kryter KD (1962) Methods for calculation of the articulation index. *J Acous Soc Am,* 34: 1689-1697.

Lawson GD and Chial MR (1982) Magnitude estimation of degraded speech quality by normal- and hearing-impaired listeners. *J Acous Soc Am*, 72 (6): 1781-1787.

Levitt H (1982) Speech discrimination ability in the hearing impaired: Spectrum considerations. *The Vanderbilt Hearing Aid Report*, Upper Darby Pennsylvania, Monographs in Contemporary Audiology; 32-43.

Licklider JCR (1944) *Report OSRD No. 4217*, Psycho-Acoustic Laboratories, Publications Board, U.S. Department of Commerce, Washington D.C.

Licklider JCR (1946) Effects of amplitude distortion upon the intelligibility of speech. *J Acous Soc Am*, 18 (2): 429-434.

Licklider JCR and Pollack I. (1948) Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech. *J Acous Soc Am*, 20 (1): 42-51.

Matsumoto J (1983) A study of threshold and loudness summation in listeners with sensorineural hearing impairments. Unpublished independent study. St. Louis, Central Institute for the Deaf.

Merklein RA (1981) A short speech perception test for severely and profoundly deaf children. *Volta Review*, 83: 36-45.

Miller GA (1947) The masking of speech. *Psychol Bull*, 44: 105-129.

Pascoe DP, Miller JD, Skinner MW, Albee D, Freiert D, and Hack Z (1980) Evaluation of the Pascoe Hearing-Aid Selection Procedure. Central Institute for the Deaf Periodic Progress Report No. 23, 51.

Pearsons KS, Bennet RL and Fidell S (1976) *Speech levels in various environments*. Bolt, Beranek and Newman, Report #3281, prepared for Office of Resources & Development, Environmental Protection Agency, Washington, D.C.

Pickett JM, Martin ES, Johnson D, Smith SB, David Z, Willis D and Otis W (1972) On patterns of speech feature reception by deaf listeners. In G Fant (ed.) *Speech Communication Ability and Profound Deafness*. Washington, DC, AG Bell Assoc. for the Deaf.

Pollack I (1948) Effects of high pass and low pass filtering on the intelligibility of speech in noise. *J Acous Soc Am*, 20: 259-266.

Pollack I (1952) On the effect of frequency and amplitude distortion on the intelligibility of speech in noise. *J Acous Soc Am*, 24 (5): 538-540.

Pollack I and Pickett JM (1959) Intelligibility of peak-clipped speech at high noise levels. *J Acous Soc Am*, 31 (1): 14-16.

Preves DA and Newton JR (1989) The headroom problem and hearing aid performance. *Hearing J*, Oct., 42 (10): 19-26.

Skinner MW (1988) *Hearing Aid Evaluation*. Englewood Cliffs, NJ, Prentice Hall.

Skinner MW and Miller JD (1983) Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. *Audiology*, 22: 253-279.

Skinner MW, Miller JD, DeFilippo CL, Dawson JK and Popelka GR (1986) Word identification by listeners with sensorineural hearing loss using four amplification systems. In TJ Glattke, MJ Collins and LA Harker (eds.) *Sensorineural Hearing Loss: Mechanisms, Diagnosis, and Treatment*. Iowa City, University of Iowa Press.

Smith CR (1975) Residual hearing and speech production in deaf children. *J Sp Hear Res*, 18: 795-811.

Thomas IB and Niederjohn RJ (1970) The intelligibility of filtered-clipped speech in noise. *J Audio Eng Soc*, 18 (3): 299-303.

Thomas IB and Ravindran A (1974) Intelligibility enhancement of already noisy speech signals. *J Audio Eng Soc*, 22 (4): 234-236.

Thomas IB and Sparks DW (1971). Discrimination of filtered/clipped speech by hearing-impaired subjects. *J Acous Soc Am*, 49 (6) part 2: 1881-1887.

Tonndorf J (1986) Auditory nonlinearities: the role of cochlear hydromechanics. *Hear Res*, 22: 223-227.

Wathen-Dunn W and Lipke D (1958). On the power gained by clipping speech in the audio band. *J Acous Soc Am*, 30 (1): 36-40.

Wegel RL (1932) Physical data and physiology and excitation of the auditory nerve. *Ann Otol*, 41: 740-779.

Witting EG and Hughson W (1940) Inherent accuracy of a series of repeated clinical audiograms. *Laryngoscope*, 50: 259-269.