DEVELOPMENT OF TEST PROCEDURES FOR EVALUATION OF BINAURAL HEARING AIDS
A Final Report

Raymond T. Carhart, Ph. D. b
Director
Northwestern University School of Speech
Speech Annex Building, Room 41
Evanston, Illinois 60201

Lamar Young, Ph. D. c
Assistant Professor
Indiana University
Department of Speech and Hearing Sciences
Speech and Hearing Center
Bloomington, Indiana 47401

Research in the rehabilitative aspects of audiology has not been as extensive as audiological research focused on the development and use of various auditory tests for diagnostic purposes. The work described here, supported by the Veterans Administration, represents an effort to restore the balance by focusing on the communicative problems experienced by hearing-impaired veterans and others, and their rehabilitative needs.

Specifically, the overall goals of these investigations have been the development of test procedures for assessing the efficiency of binaural hearing aids, and the development of methods for increasing the efficiency of hearing aid use. This report will consist of two parts: a summary of the work completed under earlier contracts, and of the projects supported by Veterans Administration Contract No. V101 (134) P—6. Greater emphasis will be placed on the reporting of the latter projects.

---

a Based on work performed under VA Contract V101 (134) P—6.

b Dr. Carhart died October 2, 1975. A pioneer in the field of audiology, he received one of the first two Ph. D. degrees granted by the School of Speech at Northwestern University in 1936. Through his personal research as well as his effective chairmanship of consulting groups, he played a leading role over a 20-year period in bringing high quality hearing aids to thousands of veterans. (See memoriam in BPR 10-24).

c This manuscript was completed by Lamar Young, Jr., Ph. D., upon Dr. Carhart's death.
A SUMMARY OF THE WORK COMPLETED UNDER EARLIER CONTRACTS

Discrimination of Speech in Quiet and in Noise

The comparative performance of normal hearers and hearing-impaired persons, in understanding speech in quiet and in noise, has been studied. This research has demonstrated quite clearly that the speech understanding of persons with sensorineural loss is disrupted to a substantially greater degree, in noise, than is speech understanding for normal hearers in comparable listening situations. This general observation — of excessive breakdown of speech discrimination in noise for persons with sensorineural hearing loss — holds true regardless of the type of background noise used, e.g., competing speech, speech spectrum noise or amplitude modulated white noise.

In all conditions it is as though the masking produced by the noise is greater for persons with sensorineural hearing loss than it is for normal-hearing individuals — or for persons with conductive hearing impairments. (As might be expected, persons with conductive hearing loss can function about as well as normal hearers in like listening situations [see Tillman, Carhart and Olsen, 1970].)

The situation is generally not improved with hearing aid use. In fact, the disruption in speech understanding in noise seems to be greater with many contemporary hearing aids. While speech intelligibility scores obtained with a hearing aid in quiet often equal speech discrimination scores achieved with high fidelity amplification, performance with the hearing aid in noise is commonly much poorer (Olsen and Tillman, 1968). This statement applies not only to test results obtained for persons having sensorineural hearing deficits, but also for individuals having conductive hearing losses — and even for normal hearers listening to hearing aid reproduction of speech in noise.

The significance of these observations is quite clear. First, when dealing with a sensorineural hearing disorder, the extent of the communication handicap cannot be specified simply on the basis of two traditional measures of hearing; namely, measurement of loss in threshold sensitivity and assessment of speech discrimination in quiet. In addition to these measures, the excessive disruption of speech understanding in the presence of other competing speech or other noise, must also be established. These measures will serve as a better indicator of the communication problems that plague the sensorineural hearing loss patient in the complex listening environments of everyday life (Carhart and Tillman, 1972). Consequently, it is important that the patient (regardless of whether he has a conductive or sensorineural hearing impairment) be tested in noise when assisting him in the selection of a hearing aid. Work
completed with the support of the Veterans Administration has shown that it is only in more difficult listening conditions, such as a competing message situation or in other background noise conditions, that differences in performance with various hearing aids can be demonstrated (see Tillman, Carhart and Olsen; Carhart and Tillman, 1970, and Carhart and Tillman, 1972).

Such information is vital, not only in terms of hearing aid selection but also for counseling the patient. It seems essential for the clinician to have information regarding the patient's aided performance in quiet and in noise if realistic levels of expectation regarding hearing aid use are to be set for him. The patient himself must then be made aware of these levels of expectation. It is thus that insightful counseling becomes of vital importance in any aural rehabilitation program.

**Head Shadow Effect**

Another finding which has emerged from this work is the role of the head-shadow effect in unilateral hearing loss cases and in monaural hearing aid use. It has been demonstrated that the magnitude of the head-shadow effect is about 6.4 dB when measured via shift in spondee threshold (Carhart, 1969). This effect can cumulate to about 13 dB in noisy environments for the unilateral hearing loss case or for the monaural hearing aid wearer, and is dependent upon whether the signal to which he wishes to attend originates from the same side as the “good” ear (aided ear) or from the opposite side. When the signal of interest comes directly to the good ear (or the ear-level hearing aid) the noise originating from the opposite side is reduced about 6.4 dB by the head shadow effect in reaching the ear (or hearing aid microphone) and consequently a 6.4 dB more favorable signal-to-noise ratio is available to the listener. However, when noise and signal locations are reversed so that the speech reaches his good ear (or hearing aid) indirectly, the signal-to-noise ratio is 6.4 dB poorer at the listener’s good (or aided) ear. Consequently, the cumulative effect of the head shadow is very substantial indeed, on the order of 13 dB (Olsen and Tillman, 1968, and Carhart, 1969).

Obviation of the head shadow effect is the primary advantage of the ear-level binaural hearing aids. The individual with bilateral hearing loss (both ears “aidable” with ear-level instruments) wearing an ear-level aid at each ear will always have one aid favorably oriented with respect to the signal of interest to him at the moment; therefore he will not be faced with the situation in which the signal-to-noise ratio can be made quite unfavorable by a simple relocation of the speech-of-interest from one side of him to the other, as is the case for the monaural hearing aid user. (For a more complete discussion of the head shadow effect, see Olsen and Carhart, 1967.)
Binaural Squelch Effect

Our test results have also confirmed a binaural squelch effect. This effect is revealed in a modest enhancement in speech understanding when a normal individual is permitted to listen with both ears in a sound field test condition in noise, as opposed to being forced to listen with only one ear, even if the monaural ear is favorably oriented relative to the origin of the speech-of-interest in the noise background. In other words, speech discrimination scores are slightly better when listening binaurally than when listening monaurally in the same conditions.

This slight enhancement in performance in binaural listening has come to be labelled as a binaural squelch effect because it is as if binaural listening results in speech discrimination scores that are better by as much as they would be if the noise were reduced or “squelched” about 3 dB (Olsen and Carhart, 1967). This binaural squelching effect can also accrue, in some instances, to some hearing impaired persons wearing binaural hearing aids.

The findings reported here, under the headings of “Head-Shadow Effect” and “Binaural Squelch Effect” have helped to lead to the development of new types of hearing aids. Specifically, the CROS (Contralateral Routing of Signals) and BICROS hearing aids are spin-offs, with other support, of that work done for the Veterans Administration. The CROS instrument has been found to be very beneficial to some persons with unilateral hearing loss; i.e., very poor or no hearing in one ear and nearly normal hearing in the other ear. Very real problems encountered by unilateral hearing loss individuals—difficulties which can be understood in terms of the head shadow effect—have been demonstrated. The CROS hearing aid solution to this problem is to place the hearing aid microphone on the side of the severely impaired ear, slightly amplify the sound picked up by this microphone and deliver it to the opposite ear via a receiver and plastic tubing held in the ear canal by an open earmold. The open earmold is essentially a skeleton of an earmold serving only to hold the tubing in place: it does not occlude the ear canal. The net effect is to overcome the head shadow effect because signals originating from the side of the impaired ear are heard via a hearing aid with its signal routed to the contralateral better ear. Of course, speech originating on the side of the good ear reaches it directly, since the ear canal is not occluded by the open ear mold.

Another benefit derived from clinical experience and application of the CROS type hearing aids is the capability of fitting persons having sharp high frequency dropoffs in their hearing threshold curves. This added benefit accrues from the fact that, with the CROS type hearing aid, only a short tubing or skeleton (open) earmold is placed in the ear canal. Under these conditions, it is only the high frequencies that are efficiently transmitted to the eardrum. Thus the high frequency em-
phasis necessary to compensate the high frequency hearing loss is achieved with a CROS type hearing aid and the open earmold. This arrangement is more effective and efficient than attempting to modify the frequency response of the hearing aid in other ways. Furthermore, the use of the open earmold allows airborne sounds to reach the tympanic membrane normally for “natural” hearing of the lower frequencies. Prior to this innovation of the CROS hearing aid and its application to sharp high frequency hearing losses, high-frequency-emphasis hearing aids were utilized, but generally it was not possible to achieve sufficient suppression of low frequency amplification, (or adequate high frequency amplification) to be satisfactory for persons having high frequency hearing losses represented by sharply sloping audiograms. Furthermore, closed earmolds were employed with these instruments, so that “natural hearing” of the lower frequency sounds was essentially eliminated. Thus, the CROS type hearing aid and the open earmold used with it have allowed successful use of hearing aids by persons having particular patterns of high frequency hearing loss who previously did not benefit from hearing aid amplification.

This approach to fitting hearing aids for high frequency hearing loss cases is particularly applicable to a significant portion of the veteran population. (The noise levels frequently encountered in various military activities are sufficiently intense to cause high frequency hearing loss.)

Another recent development currently being tested with CROS hearing aids is an adaptation labelled “focal CROS.” In this arrangement the miniature hearing aid microphone is placed in the ear canal of the individual, rather than above the ear in the hearing aid shell. This is another important development because, due to head diffraction and ear canal resonance effects, the signal levels at the entrance to the ear canal are quite different from those above or behind the pinna where the hearing aid microphones are usually located in conventional over-the-ear hearing aids or ordinary CROS type hearing aids.

The BICROS hearing aid has been found useful in cases of handicapping bilateral hearing loss, where one ear has been called “aidable” and other ear “unaidable” with present conventional aids. In such instances an ear-level aid is fitted to the aidable ear in the conventional manner, but a microphone is also placed at ear level on the other side, the side of the unaidable ear. The output from the second microphone is also delivered to the hearing aid at the aidable ear. Thus, the listener has one hearing aid and is hearing via one ear—but from two microphones. The result is that the head shadow effect is obviated by the use of two ear-level microphones even though the listener is hearing via only one ear.

For a more complete description of the CROS and BICROS hearing aids, see Harford and Barry, 1965; Harford and Musket, 1964; and Harford and Dodds, 1966.
Head Diffraction Effects on Ear-Level Hearing Aids

The Veterans Administration has supported an experiment which investigated head baffle and head shadow effects for a front-oriented and back-oriented microphone in a hearing aid casing when worn by human subjects, and also when placed on a dummy head. The results demonstrated that greater head baffle effects are observed at the front microphone than at the back microphone while the reverse is true in terms of head shadow effects (Olsen and Carhart, 1975). Results obtained with the hearing aid mounted on a dummy head were similar in some respects to those observed when the hearing aid was worn by six subjects while they differed in other respects. Comparison of the data obtained in this study with that of Wiener (1947) indicates smaller head baffle effects but larger head shadow effects at the hearing aid microphone than at the ear canal entrance. Finally, our results suggest that the reproduction of frequencies above 2000 Hz, and also a resonance peak at about 3000 Hz in the frequency response of a hearing aid, are beneficial to a hearing aid wearer.

WORK COMPLETED UNDER CURRENT CONTRACT

In the first section of this report, work supported by the Veterans Administration prior to the current contract has been summarized. The section that follows is drawn from studies that were supported by Contract V101 (134) P-6. The goals in the initial stages of this contract continued to be oriented toward finding ways to assess the efficiency of binaural hearing aids, but in the latter portions of the contract attention has been more directed toward increasing the efficiency of hearing aid selection and use.

Work that has already been reported in progress reports will be presented in a more summarized form; experiments and results that have been completed since the last progress report will be presented in more detailed form.

The Effects of Peak Clipping on Speech Intelligibility

One of the problems facing some persons with sensorineural hearing impairment is that of “recruitment,” or a reduced dynamic range between thresholds of hearing and of pain. In attempting to wear a hearing aid at a gain setting sufficient to improve speech discrimination, persons with recruitment often complain of discomfort created by the loudness of the transient peaks of the aided signal. Consequently, it would seem desirable to find a way of amplifying the speech signal, yet limiting the large moment-to-moment fluctuations in intensity which

---

3 Progress reports on these studies appeared in BPR 10-22, 10-23, 10-24, and 10-25.
are characteristic of speech. In other words, if speech could be processed such that the overall level was amplified but the peaks of speech limited, then it might benefit those persons with a hearing loss accompanied by recruitment. One well-known method of limiting large speech peaks is peak clipping.

The use of peak-clipped speech with normal listeners was actively investigated in the late 1940's and early 1950's, with the orientation of the experiments being largely toward the use of peak clipping in radio communication. Licklider (1944 and 1946) found that severe peak clipping can be imposed on a speech signal in quiet without a substantial loss in intelligibility. Licklider, Bindra, and Pollack (1948) found that speech discrimination was not impaired when the test items were infinitely peak-clipped.

Pollack (1952) mixed unclipped noise with clipped speech and found that at high signal-to-noise ratios, the intelligibility of the unclipped speech signal was greater than that of the severely peak-clipped signal; however, for low signal-to-noise ratios, the intelligibility of the latter was considerably greater than that of the unclipped signal. Pollack also investigated the effects on speech intelligibility of first filtering the speech signal and then clipping it. He found subject performance at low signal-to-noise ratios to be better when the signal was high-pass-filtered and then clipped, versus when it was either unclipped or low-pass-filtered and then clipped.

Thomas and Niederjohn (1968 and 1970) assumed that the first formant in the speech spectrum, containing substantial energy, contributed only modestly to speech intelligibility, and as a result suppressed it by first high-pass-filtering their speech signal. They then employed infinite peak clipping. In their experiment, the intelligibility of the speech was highest when the filter had a cut-off frequency of 1100 Hz and a slope of 12 dB/octave. Under these conditions, the intelligibility of the filtered/clipped speech was higher than that of unmodified speech at comparable sound levels. Thomas and Sparks (1971) investigated the use of high-pass filtering and peak clipping with hearing-impaired subjects. These investigators high-pass-filtered speech ($f_{co} = 1100$ Hz) and then infinitely-peak-clipped it. The test materials processed in this manner were then presented to 16 subjects (17 ears) having a wide variety of hearing impairment as judged by pure tone audiometry. When they were presented with filtered/clipped speech, 13 of the 17 ears showed substantial improvement in speech intelligibility over unmodified speech at all sensation levels tested.

It is possible that filtering and peak clipping might be of benefit to persons having hearing losses characterized by recruitment. Peak clipping would serve as an effective limiter of the large moment-to-moment fluctuations in the intensity of speech. Specifically, if these large fluctua-
tions in the intensity of speech—from high-energy vowels to low-energy consonants—which are so necessary for intelligibility—could be limited without decreasing intelligibility, then this should benefit those persons who have a hearing loss accompanied by a limited dynamic range. With the support of the Veterans Administration, two experiments were completed, aimed at determining the effects of filtering and peak clipping on the intelligibility of speech for normal hearers and for persons with sensorineural hearing impairment.

In the experiments reported here, speech was filtered and clipped in the following manner.

The first step was to filter speech in such a way that the spectrum of the filtered speech could be considered flat or "white." This is shown schematically in Figure 1, where the curve designated A represents an approximation of the long-term spectral characteristics of a five-talker complex (babble of voices) as determined in the laboratory. The main difference between the curve A shown here and the actual spectral content of speech is that there was a decrease in energy below 250 Hz for the five-talker complex, whereas it is assumed the energy is constant below 250 Hz. The speech stimulus was whitened by passing it through a multifilter (General Radio, Model No. 1925) which was set to the reciprocal of the speech spectrum, so that in effect, the high frequencies were emphasized. This filter characteristic is shown as Curve B. In this manner, the resulting spectrum of speech passing through the multifilter was flattened or "whitened" (see Curve C). (The rationale for whitening will be discussed further in the section describing the first study involving clipping.) By shaping the speech signal in this manner, each spectral

![Figure 1](image.png)

**Figure 1.** — Representation of the manner in which the speech stimulus was flattened or "whitened." Curve A represents the average spectrum of a five-talker complex (babble). Curve B is the reciprocal of Curve A and represents the settings of the multifilter through which the speech was passed. Line C represents the whitened spectrum of speech after being shaped by the multifilter.
component of the speech had an equal chance of reaching the criterion voltage above which clipping would occur.

Peak clipping was accomplished with a clipper constructed in the laboratories. The clipper was characterized by a (fixed) criterion input voltage and any signal exceeding this voltage was clipped. Consequently, the degree of clipping was set by varying the level of the input signal to the clipper. For example, 10 dB (decibels) of clipping could be achieved by increasing the level of the signal into the clipper 10 dB above the criterion voltage, and 20 dB of clipping required 20 dB increase in the level of the input signal above the criterion voltage. The magnitude of clipping for speech was defined in terms of a 1000-Hz calibration tone recorded on the magnetic tape at the same level as the frequent peaks of the speech signal on that tape. The intensity of the calibration tone was adjusted so that it was just below the criterion clipping voltage. This point represented minimal clipping, since just the peaks of speech above the level of the calibration tone were clipped. Ten dB of clipping for speech was obtained by increasing the level of the input speech signal by 10 dB above the criterion voltage, and 20 dB, 30 dB, 40 dB of clipping for speech were defined similarly. In addition, a 20-kHz sinusoid was electrically mixed with the speech signal at a level 6 dB greater than the intensity of the background noise on the tape. This was done in order to prevent the tape noise from intruding into the silent periods between test words for the clipping conditions. The speech signals were passed through a 10-kHz low-pass filter before presentation to the subjects, thus removing the 20-kHz tone.

Experiment 1.

Although it is clear that filtering and peak clipping might be of benefit to persons having hearing losses characterized by recruitment, the manner in which speech should be filtered and clipped to yield optimum performance is not evident. There remain several aspects regarding the high-pass-filtering and peak clipping of speech which should be investigated with normal listeners. For example, Thomas and Sparks (1971) cite experimental evidence by Martin and Pickett (1970) that low frequency energy in the first-formant in speech (F1) causes substantial masking of the higher frequency second-formant (F2) transitions for hearing-impaired listeners. Since the F2 transitions are a major cue for most consonants, and are probably the most important carriers of linguistic information in the speech signal (Liberman et al., 1967), Thomas and Sparks suggest high-pass-filtering the speech above 1100 Hz to remove the energy at F1. Such filtering, however, removes first-formant information (such as voicing and manner) from the speech signal.

An alternate approach to the problem (of possible F1 masking of F2 transitions) was not to filter out the low frequencies, but rather to
emphasize the higher frequencies. Specifically, it was of interest to
determine whether speech which had been shaped to give the low and
high frequency components equal energy was as intelligible as unmod-
ified speech. If so, could such speech then be clipped and still retain
intelligibility? Another point of interest was concerned with the fact that
the earlier studies employed only infinite peak clipping. What about the
effects on speech intelligibility of lesser degrees of peak clipping? Con-
sequently, the first investigation involving clipping sought to answer two
questions:
1. Is speech intelligibility altered when speech is filtered so that each
spectral component has equal energy (i.e., whitened)?
2. To what extent is speech intelligibility affected by different de-
grees of peak clipping ranging from minimal to infinite clipping?

Twenty normal hearing young adults served as listeners in this first
experiment. Two spondee thresholds were determined for each subject.
One threshold was obtained using unmodified spondees, and the other
employed spondees which were whitened in the manner described
earlier. (Curve B, Fig. 1.) Both speech reception thresholds were ob-
tained using the method outlined by Tillman and Olsen (1973). Articu-
lation functions were then obtained using monosyllabic words of the
CNC (consonant-nucleus-consonant) variety (Northwestern University
Auditory Test No. 6) under the following six types of speech processing:

a. Unmodified (not whitened and not clipped);
b. Whitened but not clipped; and,
c. Whitened and peak clipped to the following four degrees —
minimal, 20 dB, 30 dB, and 40 dB. These degrees of clipping were
chosen since they represented a range from minimal to infinite.

The unmodified test words were presented at sensation levels of 4, 12,
20, and 28 decibels in reference to the unmodified speech reception
threshold. In the other conditions, the test words were presented at
these same sensation levels but in reference to the filtered spondee
threshold. Thus, there were 24 experimental conditions: six modes of
speech processing times four sensation levels. The order in which the
subjects received the six modes of speech processing was randomly
determined.

Order of Presentation (Rationale)

In reference to the presentation levels, for each type of speech proces-
sing, the lowest sensation level was always used initially, followed by the
next lowest (and so forth) until the highest sensation level was utilized.
The rationale for always going successively from lowest to highest sensa-
tion level for each of the types of speech processing was to obtain the
discrimination functions with as little contamination from practice and
learning as possible. Since the first list was presented at the lowest
sensation level, only the more audible items were understandable. As successive lists were given at the higher presentation levels, the fact that items had been perceived correctly on previous presentations had no effect because items heard correctly on one level were also understood at higher levels. New words became understandable at each new presentation level and the success with these latter words was responsible for the improvement in the discrimination score.

The transducer was a TDH 39 (Telephonics Corp., Huntington, N.Y.) earphone seated in a MX41AR cushion. Subjects were tested in a double-walled test chamber (IAC 1200 Series, Industrial Acoustics Co., Inc.) with the experimenter seated in an adjacent room. Finally, the subjects were instructed to repeat each word that they heard and to guess if they were unsure. The experimenter was able to hear the subjects' verbal responses by way of earphones through a talk-back system.

Of the twenty young adults selected as subjects, seven were males and

![Figure 2](image_url)

**Figure 2.**—The discrimination functions yielded by the six types of signal processing for the normal hearing subjects. The abscissa is sensation level in decibels (dB) re the speech reception threshold.
the average age was 20.4 years. None had a threshold poorer than 15 dB HL (hearing loss) (re ANSI 1969 standards) for octave frequencies from 125 to 8000 Hz. All of the subjects presented negative histories of otological pathology.

The results yielded in this experiment can be summarized as follows. The speech reception threshold obtained with the unmodified stimuli was 18.0 dB SPL (sound pressure level) and the threshold yielded by the whitened speech was 16.0 dB SPL. The effects of whitening and peak clipping on speech intelligibility appear in Figure 2, where the percentage of discrimination words correctly repeated by the subject is shown, as a function of sensation level, for the six types of signal processing. The values shown in this figure are mean values. The standard deviations for +4 dB SL (sensation level) were approximately 16 percent for the different signals and decreased steadily as sensation level was increased. The standard deviations were only about 2 percent at 28 dB SL.

Consider first that the pattern of mean results is essentially the same for all six types of speech processing. Specifically, the six modes of speech processing yielded discrimination scores that increased linearly with increases in intensity until the sensation level was greater than 12 dB SL. At signal strengths above this magnitude, discrimination scores increased less and less, finally reaching an asymptote characterized by almost perfect discrimination at 28 dB SL.

The results yielded by the unmodified speech are very similar to those reported by Tillman and Carhart (1966) for the Northwestern University Auditory Test No. 6. For example, the normal-hearing subjects used by Tillman and Carhart reached a discrimination score of 50 percent at +4 dB SL. This is close to the performance of the subjects here, who achieved a score of 56.2 percent at +4 dB SL. Also, the subjects employed by Tillman and Carhart obtained a score of approximately 95 percent for a sensation level of +20 dB SL. This compares favorably with the discrimination score of 92.6 percent yielded by the subjects here at this same sensation level.

The differences between the results obtained with whitened and with whitened/clipped speech can be compared to the results achieved with the unmodified speech. The largest differences occurred at the +4 dB sensation level: here the unmodified speech produced the lowest discrimination score (56.2 percent) and the whitened but unclipped speech produced the highest score (67.3 percent). This is a difference of 11.1 percent and is statistically significant at the 0.05 level of confidence (ANOVA and Newman-Keuls range test).

The whitened/clipped speech at +4 dB SL produced discrimination scores that were intermediate between those yielded by the unmodified speech and those obtained with the whitened speech.

As sensation level was increased, the differences among the mean
scores became systematically smaller. For the +12-dB and +20-dB SL conditions, the unmodified speech continued to yield the lowest discrimination score, although none of the differences between the scores yielded by the various methods of speech processing at these two sensation levels was statistically significant. For the highest sensation level employed in this study (+28 dB SL), the maximum difference was between the score produced by the unmodified speech (99.1 percent) versus that yielded by whitening and 30-dB clipping. This difference is only 1.9 percent and is not statistically significant.

Consider now that the unmodified test words were presented at a sensation level which was referenced to the speech-reception-threshold determined with unmodified spondees — whereas the whitened and whitened/clipped words were presented at a sensation level re the speech-reception-threshold obtained with whitened spondees. Since the two speech-reception-thresholds differed by 2.0 dB (with the whitened

![Discrimination functions](image)

**FIGURE 3.**—The discrimination functions yielded by the six types of speech-processing for the normal-hearing subjects. These are the data from Figure 2 here plotted in terms of dB sound pressure level (SPL).
spondee threshold being lower than the unmodified speech threshold), the unmodified discrimination words were presented at an absolute sound pressure level that was 2 dB greater than the level at which the whitened and whitened/clipped speech was presented. The values shown in Figure 2 therefore represent scores that were obtained at different sound pressure levels.

In order to demonstrate the effects of the different types of speech processing for the absolute presentation levels employed in this study, the data from Figure 2 were replotted as a function of sound pressure level. This is shown in Figure 3, where the abscissa represents the presentation level in decibels (dB) re 0.0002 dyne/cm², and the ordinate represents the discrimination scores for the six different types of signal processing. Here it is evident that the whitened and whitened/clipped speech is more intelligible than the unmodified speech, especially at the lower presentation levels. For example, at 28 dB SPL, the whitened and minimal clipped speech yielded a discrimination score of 92.0 percent, whereas we would expect (by extrapolation) that a discrimination score of about 78.0 percent would be obtained with the unmodified speech.

Even at higher presentation levels, there is a trend for the whitened and whitened/clipped speech to produce higher discrimination scores. This would suggest that such signal processing slightly enhances the intelligibility of speech even under favorable conditions.

The results shown in Figure 2 and Figure 3 suggest that, under many clinically realistic conditions, peak clipping might benefit those persons who have a hearing loss characterized by a reduced dynamic range. Specifically, when the test items were presented at 28 dB SL, there was no real difference between the discrimination scores yielded by the unmodified speech and the whitened/clipped speech. At the +4-dB sensation level, there was a statistically significant (p ≥ 0.05) difference between the scores produced by the unmodified and the whitened speech, but it was in the direction of the whitened speech yielding higher scores than the unmodified speech. It is interesting that in this, the most difficult listening condition, the subjects achieved the lowest discrimination scores for the unmodified speech and that higher scores were yielded by all other types of signal processing. In addition, if the discrimination scores obtained in this experiment are plotted against the presentation intensity expressed in sound pressure level, it is apparent that whitening and clipping the speech enhanced the intelligibility of the target items, especially at low intensities. Overall, these data confirm those of earlier investigators (e.g.: Licklider, 1944 and 1946; Martin, 1950; and Pollack, 1952) who report that intelligibility was not decreased when speech was peak clipped.

An additional finding to emerge from this first study is that there was relatively little difference in speech discrimination for the various de-
degrees of peak clipping, although there was a tendency for the minimal peak clipping to yield slightly higher discrimination scores than 20, 30, and 40 dB of clipping. For example, at the +4-dB sensation level, the minimal clipped speech yielded a discrimination score of 66.9 percent compared to the discrimination scores of 59.3 percent, 61.7 percent and 60.7 percent obtained with 20, 30, and 40 dB of clipping, respectively. This would indicate that, for the 40-dB-of-peak-clipping condition (i.e., infinite clipping), where only axis-crossing information is retained, the intelligibility of speech is comparable to conditions of minimal clipping where only the large peaks of speech are clipped. Consequently, if peak clipping is to be used for limiting the dynamic range of speech, then maximum clipping apparently may be utilized without altering the intelligibility of speech to a statistically greater extent than would be accomplished by minimal clipping.

*Retaining Linguistic Information*

Another finding which can be demonstrated from these data concerns the subjects' performance with the whitened speech. Recall that Thomas and Sparks (1971), in their experiment involving clipped speech, suggested that the speech be high-pass-filtered at 1100 Hz prior to being clipped. This suggestion was based on experimental work by Martin and Pickett (1970) who found in hearing-impaired listeners that low-frequency first-formant energy masked second-formant transitions which were a major cue for consonants and served as a carrier of linguistic information. High-pass filtering at 1100 Hz, however, removes first-formant information which may be important to the listener. Consequently, in the present study first-formant information is retained by not filtering the low frequencies, while the higher frequencies are emphasized, to minimize the masking of second-formant transitions.

The speech signal was shaped in such a manner that (on a long term basis) each spectral component of the speech had equal energy, i.e., the speech signal was whitened. The discrimination functions in Figure 2 demonstrate that unmodified speech and speech which has been whitened are equally intelligible at high intensities — and whitened speech appears to be more intelligible than unmodified speech at low intensities (Fig. 3).

*Experiment 2*

In the first experiment, it was demonstrated that whitened and clipped speech remains intelligible when the stimulus is presented in quiet and when subjects have normal hearing. The purpose of this second experiment was to investigate the intelligibility of whitened and clipped speech in the presence of a competing message.

Three types of signal processing were used: unmodified, whitened,
and whitened plus 30 dB of peak-clipping.

Persons with normal hearing and also persons with sensorineural hearing loss were employed as subjects.

The general experimental procedure was to mix with the target words, electrically, a competing “message” composed of five talkers each reading a separate passage. The target words were presented at a constant level while the intensity of the competing message was varied to yield five signal-to-competition ratios. A simplified diagram of the apparatus used to accomplish this is shown in Figure 4. The test discrimination words were stored on channel 1 of a two-channel tape recorder and the competition on channel 2. The two outputs of the tape recorder went to separate attenuators and then into a summing amplifier having unity gain. The mixing of the target signal and the competition, therefore, occurred prior to whitening and clipping.

In other words, the signal going to the filter and clipper was composed of the target message and the five-talker competition — it was this composite signal which was first whitened and clipped and then presented to the subjects. To vary the signal-to-competition ratio, it was necessary to change only the attenuator controlling the competing message (labeled ATT 2 in Fig. 4). The overall level of the composite signal was determined at the amplifier-attenuator complex. The presentation level was set by grounding channel 2 of the tape recorder (the competing message) through a 600 Ω resistor and setting the level of the target material to a nominal level of 85 dB SPL. The appropriate signal-to-competition ratio was obtained by ungrounding switch S1 and selecting the necessary amount of attenuation for the competition.

The experimental design for this study included the three types of
signal processing, each presented at five signal-to-competition ratios. The modes of signal processing were unmodified, whitened, and whitened plus 30 dB of clipping.

The five signal-to-competition ratios were $-12$, $-8$, 0, 8 and 12 dB for the listeners with normal hearing.

For the hearing-impaired subjects, four signal-to-competition ratios were used: $-6$, 0, 8 and 12 dB.

The target discrimination words employed in this study were the ten Lehiste-Peterson (1962) word lists instead of the four lists of the NU No. 6 words used in the first study. This change was made so as to use the greater number of lists from Lehiste-Peterson, thereby minimizing learning effects.

Since the orientation of the experiment reported here was toward using peak clipping as a method of limiting the dynamic range of speech for the hearing-impaired, it was felt that a hearing aid receiver would be more appropriate as the final transducer. Consequently, a hearing aid receiver was employed in order to reflect more accurately the performance that might be achieved via an actual hearing aid. The receiver (a Knowles Electronics, Model No. BP 1710 adjusted to zero bias) was used with #13 tubing (#13 tubing has an i.d. of 1.9 mm) and was terminated in an EAR plug inserted in the subject's ear. The EAR plug is a self-adjusting plug of foamed polymer which expands to full ear canal diameter. A hole was drilled through the plug so that the #13 tubing could just be inserted through. The frequency response of the receiver, tubing and EAR plug is shown in Figure 5. The response was determined by fitting a KEMAR manikin (Burkhard and Sachs, 1975) with the receiver, tubing, and EAR plug, then sweeping a pure tone of constant voltage

![Figure 5](image-url)
across the receiver. The response shown in Figure 5 represents the output of the condenser microphone terminating KEMAR’s ear canal.

Two groups of subjects were employed in this second experiment. One group consisted of ten young adults: four were males and the mean age of the group was 20.4 years. These subjects had no history of otological pathology and had pure tone thresholds no poorer than 15 dB HL (re ANSI 1969 standards) for octave frequencies from 125 to 8000 Hz. The second group of subjects were eight persons with presbycusic hearing loss meeting the following criteria: first report of hearing loss at 60 years or older; speech reception threshold in better ear between 20 and 45 dB HL (re ANSI 1969 standards); and discrimination greater than 70 percent in the better ear. Five of these persons were male and the mean age of the group was 76.1 years.

Figure 6 presents the results obtained with the normal hearing sub-

![Graph](image)

**Figure 6.** The discrimination functions yielded by the three types of speech processing for the normal-hearing subjects in the presence of competition. The abscissa is signal-to-competition ratio.
jects. The abscissa is the signal-to-competition ratio expressed in decibels (dB) and the ordinate is the percentage of test words correctly repeated. The open circles represent the data obtained with the unmodified speech; the open squares are results achieved with the whitened speech and the closed circles represent the data yielded by the whitened/clipped speech. The values shown here are mean values. The standard deviations associated with these means ranged from approximately 1 percent at the −12-dB ratio (where all are low) to 10 percent to 12 percent at the 0-dB ratio.

Note that in Experiment 1, the mean discrimination scores were obtained as a function of sensation level, whereas here they were obtained by varying the signal-to-competition ratio (S/C). Zero decibels (dB) S/C is the most confusing condition; +12 dB S/C is the most audible.

The three types of speech processing yielded essentially the same discrimination functions with essentially the same shape. When the level of the competing message was substantially greater than the level of the speech (the −12-dB signal-to-competition ratio), the subjects were unable to repeat any of the test material. As the signal-to-competition ratios were made more favorable the subjects' performance increased in a linear fashion, and at very favorable signal-to-competition ratios the scores of the subjects approached an asymptotic value, after which increases in intensity did not bring about increases in discrimination. This pattern of results, seen for each of the three types of speech processing, is typical of speech testing.

In addition, two findings emerge if we compare the results yielded by the unmodified speech to those obtained with the whitened and whitened/clipped speech. First, the discrimination functions produced by the unmodified speech and the whitened stimuli are essentially identical. It can be seen in Figure 6 that the functions for these two signal processing types intertwine with each other. Secondly, the discrimination function obtained with the whitened/clipped speech is substantially different from that yielded by both the unmodified and the whitened speech. This is especially true for the more favorable signal-to-competition ratios, that is, the 0-dB, +8-dB and +12-dB ratios. At these signal-to-competition ratios, the whitened/clipped speech produced discrimination scores that were significantly lower (p ≤ 0.01, ANOVA and Newman-Keuls Range Test) than those yielded by the unmodified and the whitened speech material. Specifically, at the 0-dB signal-to-competition ratio, the unmodified speech and the whitened speech produced discrimination scores of 69.0 percent and 68.8 percent, respectively—at this same S/C ratio, the whitened/clipped speech yielded a score of only 29.6 percent. Even when subject performance was approaching an asymptotic value (the +12-dB signal-to-competition ratio) and discrimination scores of 96.6 percent and 96.2 percent were ob-
tained for the unmodified speech and the whitened speech (respectively), a score of only 79.0 percent was obtained for the whitened/clipped speech.

In other words, when a target speech signal is presented simultaneously with a competing message, the composite signal may be whitened without degrading intelligibility—but when peak clipping is imposed on this whitened signal, then a substantial reduction occurs in the intelligibility of the target speech.

The results for the listeners with presbycusic hearing loss are shown in Figure 7. Here, the data reflect the same type of trends that were seen for the normal hearing subjects. The results yielded by the unmodified speech and the whitened speech are virtually the same, whereas the scores obtained with the whitened/clipped speech are statistically (p ≤

---

**Figure 7.**—The discrimination functions yielded by the three types of speech processing for the subjects with presbycusic hearing loss in the presence of competition. The abscissa is signal-to-competition ratio.
0.01) lower than those yielded by the unmodified and whitened stimulus material. When the signal and competition had the same intensity (the 0-dB signal-to-competition ratio), the unmodified and whitened speech produced discrimination scores of 34.5 percent and 40.5 percent, respectively, while the whitened/clipped speech yielded a score of 5.5 percent at this same ratio. For the most favorable signal-to-competition ratio (+12 dB), discrimination scores of 78.0 percent and 81.0 percent were obtained with the unmodified and whitened speech respectively. These are compared to a score of only 43.5 percent yielded by the whitened clipped speech. Thus it may be concluded that, for persons with presbycusic hearing loss, peak clipping substantially reduces the intelligibility of speech relative to the intelligibility of that speech when it is unmodified or just whitened.

In addition, it may be that disruption in intelligibility created by peak clipping may be greater for persons with hearing loss than it is for listeners with normal hearing. Consider, for example, that for the normal hearing subjects, at the most favorable signal-to-competition ratio, the difference between the discrimination score obtained with the unmodified speech and that obtained with the whitened/filtered speech was 17.6 percent. In contrast, for the subjects with presbycusic hearing loss at the same signal-to-competition ratio, the difference between the two scores was 34.5 percent. Thus, one could argue that under the most optimal of the listening conditions employed in this study, the disruption created by the peak clipping was almost twice as great for the hearing-impaired subjects as it was for the listeners with normal hearing.

It is important to note that the subjects with hearing losses found the unmodified speech more difficult than did the normals. This is evidenced by the fact that at the +12 signal-to-competition ratio, the hearing-impaired subjects scored 78.0 percent for the unmodified speech versus 96.6 percent for the normal hearing subjects on the same task. This apparent poorer performance of the pathologicals may reflect the fact that for these subjects, the presentation level of the target material was not sufficiently high to be securely on the asymptotic portion of the articulation function. The mean speech reception threshold for the presbycusic subjects was 62.9 dB SPL. Recall that the presentation level of the discrimination test words was 85 dB SPL, with the level of the competition being varied around this to achieve the appropriate signal-to-competition ratios. This means that target words were presented at a mean sensation level of 22.1 dB SL. But Tillman and Carhart (1966) have shown that for hearing-impaired subjects listening to the N.U. Test No. 6, asymptotic performance is not reached until a presentation level of 40 dB SL is employed. Thus, it may be that for the +12 signal-to-competition ratio (for example), the presbycusic subjects were not on the plateau of the articulation function, therefore the
disruption created by the clipping may have lowered their discrimination proportionately more than that of the normals, who were well on the plateau.

In other words, had the presentation level of the signals been such for the hearing impaired subjects that it placed them on the same place on the plateau of the discrimination functions as the normals, then the two groups might have shown equivalent decreases in intelligibility for the peak-clipped speech. (There is obviously no reason to suspect that the clipping would have been less disruptive for the hearing impaired subjects than it was for the normal hearing listeners.)

There are two other aspects of this study which deserve comment. First, Thomas and Ravindran (1974) report data which show that peak clipping did not affect the intelligibility of a target speech mixed (before clipping) with a broad-band noise. These investigators first passed white noise through a bandpass filter having cut-off frequencies of 250 Hz and 6800 Hz and then electrically mixed this noise with the target discrimination words. The composite signal was high-pass-filtered \( (f_{co} = 1100 \text{ Hz}) \), infinitely peak-clipped, and then presented to normal-hearing subjects. Under these conditions, their listeners achieved slightly higher discrimination scores for the filtered/clipped speech than they did for the unmodified material.

The reason for the discrepancy between results reported by Thomas and Ravindran and these studies was not determined. However, it may be related to the fact that in these studies the competing message was speech and contained substantial energy below 250 Hz. In contrast, the noise used by Thomas and Ravindran was bandpass-filtered at 250 Hz and 6800 Hz and then high-pass-filtered along with the speech signal at 1100 Hz. Evidence to support this line of reasoning is found in a series of experiments conducted by Licklider (1944). Licklider mixed a broad band noise (roughly comparable in terms of spectrum to the five-talker “babble” speech competition) with the target words, and then infinitely-peak-clipped the composite signal. Under these conditions, he found a reduction in the intelligibility of the clipped speech compared to what was seen when the clipped speech was presented in quiet. He concludes that “clipping is more seriously detrimental when noise . . . is present in the signal passing through the nonlinear circuit than when the signal is kept free from noise . . .” (p. 69).

Secondly, Thomas and Sparks (1971) report an increase in speech intelligibility for sensorineurals presented with high-pass-filtered and clipped speech. Their data were obtained with one speaker, that is, in the absence of a competing message. If the +12-dB signal-to-competition ratio employed in our study was sufficiently favorable that the data obtained at this ratio should be considered roughly comparable to results obtained in quiet, then our data contradict those of Thomas and

30
Sparks—since we show a reduction in discrimination of peak-clipped speech for subjects with sensorineural hearing loss. The reason for this disparity in findings remains to be ascertained.

The present study, however, differed from the Thomas and Sparks experiment in terms of the type of filtering that was imposed on the speech before clipping.

Moreover, subject selection may have substantially influenced the results. Thomas and Sparks do not report the audiological classification of their subjects. However, they do give the audiogram for each of their listeners, and if we take those audiograms, which by visual inspection appear to be comparable to those of our presbycusic subjects, then only three of the subjects used by Thomas and Sparks should be used for comparison. These are subjects c, h and i. Of these three subjects, only one demonstrated a distinct advantage for the clipped speech. The other two subjects yielded essentially the same discrimination scores for clipped and unmodified speech at the highest sensation level tested. Also, it should be pointed out that the subjects in the Thomas and Sparks study were substantially younger than the prebycusics employed here.

Conclusions Drawn from Experiments 1 and 2.

These two experiments demonstrate that the whitening and peak clipping of speech does not degrade its intelligibility for normal-hearing persons under optimal listening conditions (i.e., in the absence of competition). However, speech intelligibility is substantially reduced when the target speech and a competing message are whitened and then peak clipped. This is true for persons with normal hearing as well as for those with hearing impairments. Because of the degradation in speech intelligibility created by peak clipping when a competing message is present, this method of signal limiting would not seem desirable in a wearable hearing aid, at least not at the present time.

The Effects of Amplitude Compression on Speech Intelligibility

Another approach to limiting the dynamic range of speech is to use amplitude compression. Since peak clipping significantly reduces speech intelligibility in the presence of competition, the problem of limiting the dynamic range of speech was approached by employing amplitude compression. The aim was still the same—to limit the large moment-to-moment fluctuation in the intensity of speech without decreasing intelligibility; if this aim could be accomplished, then it should benefit those persons having a sensorineural hearing loss characterized by a limited dynamic range.

After a discussion of the results of other investigations in the general area, there follows the report of two studies which investigated the
effects on speech intelligibility of amplitude compression under different conditions and for different listeners.

Edgarth (1952) first suggested that compression might benefit persons with certain types of hearing loss. Parker (1953) reported improved speech intelligibility with high pass filtering and compression amplification on 8 of 10 subjects having sensorineural hearing loss. Kretsinger and Young (1960) employed compression on normal listeners and found improved intelligibility in noise. Later, however, Caraway and Carhart (1967) concluded on the basis of their results that neither 2 : 1 nor 3 : 1 compression ratios\(^2\) offered any important advantage over 1 : 1 amplification when comparisons were made in terms of intelligibility at a given sensation level of the output signal. Vargo and Carhart (1972) completed additional experiments which supported the conclusions of Caraway and Carhart.

On the other hand, Burchfield (1971) reported an increase in intelligibility for the 2 : 1 and 3 : 1 ratios of compression over linear amplification. Villchur (1973) reported a substantial increase in intelligibility for six subjects having sensorineural hearing losses when two-band compression and frequency equalization was utilized. Yanick (1973) compared hearing aids having compression amplification with hearing aids employing linear amplification, and found that the amplitude compression aids produce dramatically improved discrimination scores in quiet.

It is clear, therefore, that filtering and amplitude compression might be of benefit to persons having hearing losses, especially those characterized by recruitment. However, there remain several aspects regarding amplitude compression which should be investigated. For example, several investigators have suggested some form of speech shaping which attenuates or removes low frequency energy to improve intelligibility (see Thomas and Sparks, 1971; Thomas and Pfannebecker, 1974; and Thomas and Ravindran, 1974). Parker (1953) found that when speech was first high-pass-filtered at 670 Hz and then amplitude-compressed, speech intelligibility was improved for some subjects with sensosineural hearing impairment. Consequently, it was thought that it would be of interest to shape the speech by emphasizing the high frequencies, rather than by removing low frequency information, and then impose different degrees of amplitude compression. In brief, it was planned to whiten the speech for the same reasons and in the same manner as that used in the earlier studies involving peak clipping.

In addition, if amplitude compression is to be utilized to reduce the

---

\(^2\) The term "compression ratio" refers to the gain reduction of the amplifier as determined by the input-output function. For example, a 2 : 1 compression ratio is achieved when an increase in input of \(n\) dB produces an increase in output of \(n/2\) dB (e.g. an increase in input of 10 dB results in an increase in output of 5 dB).
dynamic range of speech, it still must be determined what compression ratio should be used to obtain maximum intelligibility. For example, the majority of experiments have involved compression ratios on the order of 2:1 to 5:1 (see Yanick, 1976); but Vargo and Carhart (1974) report data which indicate that subjects achieve essentially the same discrimination scores for speech undergoing 5:1 compression as for speech undergoing 20:1 compression. Consequently, it was thought that it would be of interest to determine the intelligibility of compressed speech under compression conditions that are more commonly used (e.g., a ratio of 3:1) and also for a compression ratio that more severely limits the dynamic range of speech (e.g., a ratio of 10:1).

Finally, the effects of compression on speech intelligibility when there is a competing message are not known. Most of the studies involving compression have presented the processed speech in quiet-to-normal hearing situations, although Yanick (1975) has tested the intelligibility of compressed speech in the presence of a competing message (composed of four talkers) with subjects having sensorineural hearing losses. Yanick, however, does not report the performance of his subjects for unmodified speech — only their discrimination scores for compressed speech. Thus, it cannot be determined whether the compression reduced speech intelligibility below what could be achieved with unmodified speech — it may be that while compression does not reduce speech intelligibility for a single talker, the intelligibility of the target speech might be substantially reduced by compression when two or more talkers are present.

Therefore it seemed important to investigate whether compression reduces speech intelligibility in the presence of a competing message. Moreover, it seemed that such an investigation should utilize subjects with normal hearing and also listeners with sensorineural hearing impairment.

For these reasons, two experiments were completed involving amplitude compression. The first utilized only normal-hearing subjects. It investigated the effects of whitening the speech prior to compression, and compared the use of a moderate compression ratio (3:1) with one that reflected a more drastic reduction of the dynamic range of speech (10:1).

(The second study investigated the effects of amplitude compression on speech intelligibility when more than one speaker was present. This second study utilized both normal and hard-of-hearing subjects.)

Compression Experiment 1

Ten normal-hearing young adults served as subjects in this first experiment, the purpose of which was to determine the intelligibility of speech that was first whitened and then amplitude-compressed. The in-
Instrumentation was essentially unchanged from the clipping experiments except that the peak clipper was removed and a compression amplifier (Spectra Sonics, Model 610) was inserted in its place. Thus, the speech signal was first whitened by passing it through a multifilter set to the reciprocal of the speech spectrum and then compressed via the compression amplifier. Two compression ratios were used in this study—3 : 1 and 10 : 1 (In the first case, an increase of 3 dB in the signal to the input of the compression amplifier resulted in an increase in the output of the amplifier of 1 dB. In the latter case, an increase in the input of 10 dB resulted in an output increase of 1 dB.) Attack time of the compression system was fixed at 100 ns (nanoseconds) with recovery adjusted to 20 msec—the results of Lynn and Carhart (1963) suggest that these values result in optimum subject performance.

The experimental design for this study included four types of signal processing, each presented at five presentation levels. The four types of signal processing were: (i) unmodified; (ii) whitened; (iii) whitened with a 3 : 1 compression ratio; and (iv) whitened with a 10 : 1 compression ratio. The presentation levels were from 12 to 44 dB SPL in increments of 8 dB.

The test items were the 1.0 Lehiste-Peterson word lists used in the peak clipping experiments.

The order in which the four modes of speech processing were utilized with a particular subject was randomly determined. In reference to the intensity levels of the speech, this always went successively from the lowest presentation level to the highest presentation level for each of the types of speech processing. This was done in order to obtain the discrimination functions with as little contamination from practice and learning as possible. Finally, the hearing aid receiver from the peak clipping studies was used in order to more accurately reflect the transducer that would be used with an actual hearing aid.

The ten young adult subjects had an average age of 20.7 years; seven were males. None of the listeners had a threshold poorer than 15 dB HL (re ANSI 1969 standards) for octave frequencies from 125 to 8000 Hz. All of the subjects presented negative histories of otological pathology.

The results from this first experiment are presented in Figure 8, showing the percentage of test words correctly repeated for each of the four types of speech processing, as a function of presentation level. The values in Figure 8 are mean values with the accompanying standard deviations ranging from 2.1 percent at the 44 dB SPL level for the unmodified speech to 14.4 percent at the 28 dB SPL level for the whitened speech with 10 : 1 compression.

Consider first that the pattern of mean results is essentially the same for the four modes of speech processing. Specifically, at 12 dB SPL, the subjects were unable to perceive any of the test items for the four types of
processed speech. At signal strengths above 20 dB SPL, each of the speech stimuli yielded discrimination scores that increased apparently linearly with increases in intensity until the presentation level was greater than 36 dB SPL. For a presentation level of 44 dB SPL, it appears that the subjects have reached asymptote performance, characterized by almost perfect discrimination.

In addition, the results yielded by the unmodified speech are very similar to those reported by Tillman and Carhart (1966) for the Northwestern University Auditory Test No. 6. For example, the normal hearing subjects used by Tillman and Carhart reached a discrimination score of 50 percent at approximately 26 dB SPL. This is comparable to the performance of our subjects who achieved a score of 50 percent at 26.5 dB SPL. In addition, the listeners employed by Carhart and Tillman yielded a discrimination score of 95 percent for a signal intensity of 42

![Figure 8](image-url)  
**Figure 8.** The discrimination functions yielded by the four types of speech processing for normal hearing subjects.
dB SPL. This compares favorably with the discrimination score of about 94 percent that can be extrapolated from the discrimination function in Figure 8.

Comparing the results yielded by the four types of speech processing, it is apparent (Fig. 8) that the whitened speech yielded discrimination scores that tended to be poorer than the results obtained with the other three speech types. This poorer performance seen with the whitened speech is statistically significant \(p < 0.01\), ANOVA and Newman-Keuls Range Test) for the results at the 28 and 36 dB SPL presentation levels. Consequently, it must be concluded that whitening in this instance produced a slight degradation in speech intelligibility, although the decrement was not maintained as performance level approached asymptote. This finding contradicts the experimenters' earlier studies involving whitening and peak clipping in which whitened speech yielded comparable (and in some cases slightly higher) discrimination scores compared to results obtained with unmodified speech.

It will be noted that the unmodified speech, the whitened speech with 3:1 compression and the whitened speech with 10:1 compression yielded essentially the same results. The largest differences among the results obtained with these three types of processed speech are found at 36 dB SPL: at this intensity the largest difference is between the data obtained with the unmodified speech versus that produced by the whitened speech with 10:1 compression—and this difference is only 4.8 percent.

The fact that the discrimination scores produced by the whitened speech with both 3:1 and 10:1 compression were not significantly different from the scores obtained with the unmodified speech, indicates that amplitude compression does not degrade speech intelligibility. Also, it is important to note that compression ratios of 3:1 and 10:1 produced essentially the same subject performance, indicating that under conditions of substantial compression (the 10:1 ratio), the intelligibility of speech is comparable to conditions of significantly less compression (the 3:1 ratio). Consequently, if amplitude compression is to be used for limiting the dynamic range of speech, then it would appear that substantial compression may be utilized without altering the intelligibility of speech to a greater extent than would be accomplished by a more minimum degree of compression.

These results confirm those of earlier investigators (see Yanick, 1973 for a complete review) suggesting that intelligibility remains essentially unchanged when the speech signal undergoes amplitude compression.

Compression Experiment 2.

Since hearing-impaired listeners are often required to listen to a particular speaker when other talkers are present, it seemed important
to determine the intelligibility of whitened and compressed speech in the presence of a competing message. This experiment, therefore, investigated the effects of amplitude compression on speech intelligibility when more than one speaker is present. The four types of signal processing employed in the previous experiment were used again (unmodified, whitened, whitened with 3:1 compression, and whitened with 10:1 compression). Persons with normal hearing and also persons with sensorineural hearing loss were employed as subjects. A competing message composed of five talkers was electrically mixed with the target words. Discrimination functions were generated by presenting the target words at a constant level of 85 dB SPL and varying the intensity of the compet-

![Figure 9](image-url)  
**Figure 9.**—The discrimination functions yielded by the four types of speech processing in the presence of competition. Data are shown for the listeners with normal hearing and also for the hearing-impaired subjects. The abscissa is the signal-to-competition ratio.
ing message to yield three signal-to-competition ratios. The signal-to-competition ratios were 0, +8 and +12 dB. The target discrimination words employed in this study were the ten Lehiste-Peterson word lists used in the previous studies. Also as in the earlier study, the final transducer was the hearing aid receiver.

This experiment employed two groups of subjects. One group consisted of 10 young adults, five male and five female. The average age was 24.0 years and all met the criterion (outlined earlier) used to select subjects with normal hearing in previous studies. The second group of subjects were 10 persons with presbycusic hearing loss meeting the following criteria: first report of hearing loss at 60 years or older; speech reception threshold in better ear between 20 and 45 dB HL (re ANSI, 1969 standards); and discrimination greater than 70 percent in the better ear. Seven of these persons were male; the mean age was 77.1 years.

Figure 9 presents the results obtained with both groups of subjects. The abscissa shows the signal-to-competition ratio expressed in decibels (dB) and the ordinate displays the percentage of test words correctly repeated. The open symbols present the data obtained with the normal hearing subjects for the four types of speech processing and the closed symbols are the results produced by the hearing-impaired subjects for the same four types of signal processing. Circles represent data obtained with unmodified speech; squares are results achieved with whitened speech; triangles present scores produced by whitened speech with 3 : 1 compression; and inverted triangles present data yielded by whitened speech with 10 : 1 compression. The values shown are mean values.

Standard deviations associated with these means are on the order of 5 to 8 percent for the normal hearing subjects and 11 to 13 percent for the hearing impaired subjects.

From the data obtained (in Compression Experiment 1) with the normal-hearing listeners, it is clear (Fig. 9) that whitening the composite speech signal did not degrade the intelligibility of the target items. This is evidenced by the fact that the unmodified and the whitened speech signal produced essentially the same discrimination functions. On the other hand, the amplitude compressed speech yielded discrimination scores that were slightly lower than those produced by the unmodified speech. The largest difference was at the +8-dB signal-to-competition ratio between the unmodified speech and the whitened 3 : 1 compressed speech. This difference is 9.6 percent and is significant at \( p \leq 0.01 \). Also, at the +8 signal-to-competition ratio the results obtained with the 10 : 1 compressed speech are significantly \( (p \leq 0.05) \) lower than those yielded for the unmodified speech.

Thus, it would appear that compression amplification does lower the intelligibility of speech for normal hearers when more than one talker is
present. However, the reduction in intelligibility is not substantial, although it is statistically significant in some cases.

Turning now to the results obtained with the hearing-impaired listeners, three findings emerge from the data (Fig. 9):

1. The whitened stimulus was as intelligible as the unmodified speech — the discrimination functions for these two types of signal processing are essentially the same.

2. The 3:1 and 10:1 amplitude compression ratios did not yield the same results. The results obtained with the 10:1 compression were significantly lower (p ≤ 0.01) than those produced by the 3:1 compression, except at the +8-dB signal-to-competition ratio. Consequently, even though the 3:1 and 10:1 compression ratios yield similar results when only one talker is present, they produce substantially different data when the signal is composed of several talkers.

3. Speech intelligibility was degraded when the whitened speech was amplitude-compressed. This is true for both the 3:1 and 10:1 compression ratios. Even for the most favorable listening conditions (the +12-dB signal-to-competition ratio) the discrimination score yielded by the 3:1 compressed speech was 12.2 percent poorer than the score obtained with the unmodified speech. The 10:1 compressed speech produced a discrimination score 21.8 percent poorer than the unmodified speech. These differences are statistically significant at the 0.01 level of confidence.

Consequently, it would appear that amplitude compression interferes with the intelligibility of a speech target when more than one talker is present. Moreover, the disruption in intelligibility created by the compression may be greater for persons with hearing loss than it is for listeners with normal hearing.

For example, at the most favorable signal-to-competition ratio, there was a difference of 4.2 percent between the scores yielded by the unmodified speech and those obtained with the 10:1 compressed speech for normal hearing subjects, and this difference was 21.8 percent for the same comparison for the presbycusis subjects. This would seem to indicate that even for the most optimal listening conditions used in this study, the disruption created by the amplitude compression was substantially greater for the hearing impaired subjects than it was for the subjects with normal hearing.

Notice that the subjects with hearing losses found the unmodified speech more difficult than did the normals. This is evidenced by the fact that at the +12 signal-to-competition ratio, the hearing impaired subjects scored 78.4 percent versus 95.6 percent for the normal hearing subjects on the same task. Thus, the amplitude compression may have been far more disruptive to the hearing-impaired subjects (having discrimination difficulty under the optimal signal-to-competition ratio
tested) than to the normal-hearing listeners who had relatively good
discrimination.

However, the apparent poorer performance of the hearing-impaired
listeners may reflect the fact that the presentation level of the target
material was not sufficiently high to be on the asymptote portion of the
articulation function. The mean speech reception threshold for the
presbycusis subjects was 52.8 dB SPL; therefore, the target materials
were presented at a sensation level of approximately 22 dB SL, an
intensity well below the presentation level of +40 dB SL required by
hearing-impaired subjects to reach asymptotic performance (Tillman
and Carhart, 1966). Thus, at the presentation level of 85 dB SPL em-
ployed in this study, the hearing impaired subjects may not have been
functioning at an asymptotic discrimination level. If so, the disruption
created by the compression may have lowered their discrimination
proportionately more than it did for the normals, who were well on the
plateau.

Had the presbycusis subjects received the speech at an intensity suffi-
cient to place them on the same place on the plateau of the discrimina-
tion function as the normal-hearing subjects, then the compression
might have been no more disruptive for the hearing-impaired subjects
than it was for the normal-hearing listeners.

Conclusions Drawn from Experiments 1 and 2

These two experiments have demonstrated that whitening and
amplitude-compressing speech does not degrade its intelligibility for
normal-hearing persons when no competing signal is present. However,
when the target speech items are embedded in a competing message,
there is a slight reduction in intelligibility for listeners with normal
hearing. This reduction is speech intelligibility in the presence of com-
petition may be more severe for hearing-impaired listeners.

It is unclear, however, as to whether compression would be found to
reduce speech intelligibility more for hearing-impaired listeners than it
does for normals if the intensity of the target items were to be equated in
terms of sensation level. Here is a matter for investigation. Similarly,
Vilchur (1973) and Yanick (1975) have reported improvement in
speech discrimination when two-band compression is employed. It is
possible that such two-band compression might not produce a substan-
tial decrease in the intelligibility of the target speech in the presence of a
competing message—but this is an additional matter for research.

SUMMARY

A major concern of this work has been the gathering of experimental
data in the area of rehabilitative audiology. The specific goal of this
research has been knowledge regarding hearing aids and their use. Specifically, since the Veterans Administration issues hearing aids to so many individuals, the research focus has been, (i) on learning what potentials and limitations hearing aids hold for users in everyday life, and (ii) on developing techniques for evaluating these potentials and limitations at the time a veteran’s hearing aid is selected. The investigators’ past findings have contributed to procedures now employed by the Veterans Administration in the selection of hearing aids, to its policies governing the procurement of hearing aids, and to its clinical practices in the evaluation of hearing aids.

Specifically, previous work for the Veterans Administration has studied the speech understanding capability of normal hearers and hearing-impaired persons in quiet and in the presence of competition. A finding has been that, in general, individuals with sensorineural hearing loss experience a breakdown in speech understanding in noise which is considerably greater than that experienced by normal-hearing persons in the same noise backgrounds. In other words, moderately noisy listening situations which are not difficult at all for normal-hearing individuals can be wholly impossible for persons with sensorineural impairment. A second finding is that this disadvantageous situation is often made seriously worse by contemporary hearing aids.

The implication of these findings has clearly been that testing against a substantial sound background gives insight into the hearing problems encountered by hearing-impaired individuals in everyday situations. It is not sufficient simply to compare unaided and aided speech discrimination in quiet — tests of speech understanding must be delivered against some form of competition and must be assessed in a way that evaluates the extent to which background noise disturbs speech understanding for the hearing-impaired individual.

Another significant bit of information arising from this research for the Veterans Administration has been the role of the head-shadow effect on the reception and understanding of speech by unilateral hearing cases and monaural hearing aid users. When the speech to be understood comes from one side of the head and the competition comes from the other, a person with a unilateral hearing loss or a monaural hearing aid can suffer as much as a 13-dB disadvantage whenever the speech he wants to understand originates from the side opposite his good ear. It is little wonder that such individuals find that noisy environments, where conversation is shifting from person to person, are very disruptive.

One of the spin-offs of this line of research has been the development of new types of hearing aids. These types, now being provided for veterans through VA Audiology Clinics, are variations on the CROS principle. They take cognizance of the head-shadow effect, either com-
pensating for its disadvantages for some kinds of cases, or utilizing the peculiarities of the head-shadow effect to achieve a benefit for other kinds of cases. The arrangements include the CROS, the BICROS, the power CROS, the open CROS, and the focal CROS. Such devices have greatly enhanced the benefits of hearing aid amplification for veterans with unilateral hearing impairments, for those with need for an ear-level hearing aid having extra power, and for those with high frequency loss (such as that typically arising from exposure to intense noise).

In the VA contract just completed, research attention has been directed more toward increasing the efficiency of hearing aid selection and use. An example concerns a problem facing some persons with sensorineural hearing impairment in which there is a reduced dynamic range (recruitment). In their attempts to wear a hearing aid at a gain setting high enough to improve speech discrimination, persons with recruitment often complain of discomfort created by the loudness of the transient peaks of the aided signal. Thus, an investigation was made into ways of processing speech such that the overall level would be amplified while the troublesome peaks of speech were limited. Such processing, if incorporated into a hearing aid, would benefit those persons with a hearing loss accompanied by recruitment. Experimentation was made with two methods of limiting transient peaks of speech: peak clipping and amplitude compression.

In the first study with peak clipping, the speech signal was shaped in such a way that each spectral component had equal energy ("whitened") and then varying degrees of peak clipping were imposed. The results indicated that for persons with normal hearing listening in quiet, peak clipping does not substantially degrade speech intelligibility. In fact, when the speech is presented at a relatively low intensity, the clipped speech is slightly more intelligible than non-clipped speech. However, when the target message is mixed with a babble of other voices, then the intelligibility of the clipped material is far poorer than the intelligibility of the test words unmodified. Moreover, the decrease in intelligibility of the peak clipped material appears to be greater for persons with sensorineural hearing loss than for persons with normal hearing.

Based on these findings we would conclude that peak clipping as a method of signal limiting is probably not very helpful in a wearable hearing aid.

More encouraging were the results from the experiments involving amplitude compression. Here the procedure was to first whiten the speech and then impose two different degrees of compression (3:1 and 10:1). The data from the first study indicate that compression does not affect speech intelligibility for material presented in quiet, to normal hearers. The second study mixed a competing message with the target material and imposed compression on the composite signal. Both
normal-hearing subjects and persons with sensorineural hearing losses were used as listeners. The results demonstrated that for normal hear-
ers, compression reduces speech intelligibility only slightly, although the reduction is statistically significant. The reduction in intelligibility created by the compression is on the order of 5-to-6 percent at the most favorable signal-to-competition ratio. However, the reduction in intel-
ligibility caused by the compression may be more substantial for hearing-impaired listeners than for normal-hearing subjects. This is especially true for the 10 : 1 compression. Nonetheless, for this mode of speech processing, the reduction in speech intelligibility is less than it was for the peak clipped speech. Moreover, the reduction in intelligibility in compressed speech for hearing-impaired people which was found may reflect the influence of listening to the target speech at a lower sensation level than did the normal hearing subjects. Consequently, it may be that additional experimentation will demonstrate that amplitude compression, in a wearable hearing aid delivering adequate SPL, is a feasible way of limiting the dynamic range of speech.

These investigations supported by the Veterans Administration have added to the knowledge of the communication problems experienced by veterans with hearing losses, and they have provided data that can supply valuable insights for those who counsel people with hearing losses, through understanding of the problems they face in communica-
tion. In the later work involving peak clipping and amplitude compres-
sion ways of providing overall amplification while limiting the intensity of transient signals that are often painful to the hearing-impaired have been investigated. These have been important projects and have con-
tributed to the knowledge of rehabilitative audiology. This work has added to the ability to understand more fully the problems of the hear-
ing impaired and to ameliorate the consequences of a hearing loss.

BIBLIOGRAPHY
OF ARTICLES RESULTING FROM WORK SUPPORTED BY
THE VETERANS ADMINISTRATION

Carhart, R.: Problems of the Hearing Impaired in Noisy Social Gatherings. Oto-Rhino-


Carhart, R.: Principles of Competing Speech Audiometry. (C. Røjskjoer, Ed.) Andelsbog-
trykkerie; Odense, Denmark, pp. 112-122, 1970.

Carhart, R.: Monaural vs. Binaural Speech Perception. (C. Røjskjoer, Ed.) Andelsbog-
trykkerie; Odense, Denmark, pp. 218-225, 1970.

Carhart, R., and T. W. Tillman: Interaction of Competing Speech Signals with Hearing

REFERENCES


