HEARING AIDS—A REVIEW OF
PAST RESEARCH ON
LINEAR AMPLIFICATION,
AMPLITUDE COMPRESSION,
AND FREQUENCY LOWERING

ASHA

ASHA MONOGRAPHS

HEARING AIDS—A REVIEW OF PAST RESEARCH ON LINEAR AMPLIFICATION, AMPLITUDE COMPRESSION AND FREQUENCY LOWERING

Hearing Aids—A Review of Past Research on Linear Amplification, Amplitude Compression, and Frequency Lowering

Louis D. Braida
Nathaniel I. Durlach
Richard P. Lippmann
Bruce L. Hicks
William M. Rabinowitz
Charlotte M. Reed

Communications Biophysics Group
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

ASHA MONOGRAPHS NUMBER 19

AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION

Contents

Preface

During the early 1970s we became interested in exploring whether our experience in psychophysics, auditory perception, and engineering could be usefully applied to the problem of developing improved hearing aids. More specifically, we became interested in the development of improved signal-processing techniques to match speech to residual auditory function in listeners with sensorineural hearing impairments. Our goal was either to develop improved techniques or to determine the fundamental reasons why such techniques could not be developed.

Our initial efforts in this area consisted of attempting to acquaint ourselves with relevant previous research. This task, however, turned out to be a formidable one. In particular, we encountered the following difficulties. First, there existed hundreds of relevant articles dispersed throughout a wide variety of journals. Second, many of these articles failed to specify the conditions of the experiments adequately (for example, acoustic factors, characterization of the subject's impairments, and test materials). Third, there were often discrepancies between the actual test results and the conclusions drawn by the authors on the basis of their results. Fourth, there were no adequate review papers available. Those that existed contained (at least, for our purposes) too much selective filtering and interpretation on the part of the authors. Thus, for example, the reviews often ignored a large fraction of the studies performed and, for a given study considered, often ignored important details of the test procedures or test results.

This monograph provides a more comprehensive and detailed review of past research than is currently available. Our initial work on this review was performed in connection with a grant application to the National Institutes of Health. The reactions to this material on the part of our professional colleagues stimulated us to rewrite the material for publication. Although we realize that this monograph suffers from a number of defects (for example, it is "hard reading" and the conclusions drawn are not very incisive), we are hopeful that it will enable future investigators to have an easier time acquainting themselves with past research than we had. Since the final draft of this monograph was prepared in January 1978, readers should consult other sources for later reports.

The contributions to this monograph by the various authors listed on the title page vary considerably. In particular, the main body of work was performed by the first three authors. At the time of writing, Richard P. Lippmann, William M. Rabinowitz, and Bruce L. Hicks were graduate

ABSTRACT

This monograph provides a comprehensive, detailed review of past research on the three main forms of signal processing that have been studied to match speech to residual auditory function in listeners with sensorineural impairments: linear amplification, amplitude compression, and frequency lowering. In the area of linear amplification, the review focuses on the choice of frequency-gain characteristic. In the area of amplitude compression, although various types of compression are considered, the main concern is with syllabic compression. In the area of frequency lowering, a wide variety of schemes are discussed.

Despite the large amount of research that has been performed, it is extremely difficult to draw firm conclusions concerning the true potential of the various processing schemes studied. This is caused in part by the intrinsic complexity of the problem and in part by the inadequacies of the research. On the whole, the research has not been sufficiently careful, systematic, or analytic to advance fundamental understanding of the advantages and disadvantages of the various schemes.

Chapter I

INTRODUCTION

The most important element in the design of acoustical aids for listeners with hearing impairments is the selection of signal-processing schemes to match acoustical signals to residual auditory function. Moreover, since speech constitutes the most important class of such signals, it is essential that the schemes improve speech perception for such listeners. In this monograph we review previous work in this area, with special attention given to listeners suffering from sensorineural impairments.

Although speech perception, even for normal listeners, is not thoroughly understood, much is known about the cues used by normal listeners to decode acoustic speech signals (for example, see Studdert-Kennedy, 1976). When speech is presented to listeners with sensorineural impairments, however, many of the cues that are available to normal listeners may be lost or seriously degraded. Among the factors that may interfere with the detection of these cues are the following. First, and most obvious, is the loss in absolute sensitivity. Many elements of the speech signal may not be heard. Second, most sensorineural losses are accompanied by a reduction in auditory area without an equivalent increase in the ability to resolve signals within this area (thus implying a reduction in channel capacity). Consequently, even if the listener is aided by amplification, his ability to extract information from the speech signal is likely to be reduced. Third, the representation of signals within the auditory area is often seriously degraded. For example, there may be a reduced ability to discriminate changes in frequency, amplitude, or time; an abnormally rapid growth of loudness with stimulus intensity (recruitment); an abnormally rapid decrease of loudness with stimulus duration (tone decay); and abnormal distortion and spread of masking (in both frequency and time). In addition, there may exist internally generated sounds (tinnitus) which not only contribute to the loss in absolute sensitivity, but also distort the perception of external stimuli that are well above threshold. In general, these factors place a severe burden on the listener's ability to decode speech signals, even if there are no additional degradations of a more central nature (for example, concerning basic language competence, processing rate, and short-term memory). Although it is well known that speech perception is often seriously degraded by sensorineural impairments (even when amplification is applied) and that

1

many of the above-mentioned features may be related to this degradation, details of this degradation are not yet well understood. Results of studies concerned with the detailed speech perception errors made by listeners with sensorineural impairments are available in such articles as Siegenthaler (1949, 1954); Oyer and Doudna (1959); Rosen (1962); Kopra, Strickland, and Blosser (1967); Lawrence and Byers (1969); Cox (1969); Pickett et al. (1972); Owens, Benedict, and Schubert (1972); Bilger and Wang (1976); and Wang, Reed, and Bilger (1978). General background information on sensorineural impairments are available in such texts as Davis and Silverman (1970), Graham (1967), Jerger (1973), Katz (1972), and Wolstenholme and Knight (1970).

By "matching speech to residual auditory function" we mean selecting a signal-processing scheme that alters the speech signal in such a way that (to the extent possible) the lost cues are restored or that new cues that serve similar functions and that can be learned through training are introduced. Also, of course, the processing scheme must preserve those cues that are important to speech perception and that are not degraded by the impairment (for example, cues related to the rhythmic structure of speech). This problem of sharpening degraded cues or introducing new ones to replace those that are lost, while simultaneously preserving the useful ones, together with the problem of training the listener in the use of the new cue system, makes the matching task an extremely challenging one.

The principal forms of signal-processing to match speech to residual auditory function that have been studied in the past are linear amplification, amplitude compression, and frequency lowering. Amplification is addressed to the problem of elevated absolute threshold, amplitude compression to the problem of reduced dynamic range, and frequency lowering to the problem of negligible residual hearing at the higher frequencies. In Chapter II, we review previous work on linear amplification and, in particular, on the effects of frequency-gain characteristic. In Chapters III and IV, we review work on amplitude compression and frequency lowering. For convenience, the references cited in each chapter are presented at the end of that chapter.

REFERENCES

BILGER, R. C., and WANG, M. D., Consonant confusions in patients with sensorineural hearing loss. *J. Speech Hearing Res.*, **19**, 718-748 (1976).

COX, P. B., The identification of unfiltered and filtered consonant-vowel-consonant stimuli by sensorineural hearing impaired persons. Unpublished Ph.D. thesis, Univ. of Pittsburgh (1969).

DAVIS, H., and SILVERMAN, S. R., *Hearing and Deafness.* New York: Holt, Rinehart, and Winston (1970).

GRAHAM, A. B., *Sensorineural Hearing Processes and Disorders.* Henry Ford Hospital International Symposium (A. B. Graham, Ed.), Boston: Little Brown and Co. (1967).

JERGER, J. D., *Modern Developments in Audiology.* New York: Academic Press (1973).

KATZ, J., *Handbook of Clinical Audiology.* Baltimore: Williams and Wilkins Co. (1972).

KOPRA, L. L., STRICKLAND, L. E., and BLOSSER, D., Perceptual confusion analyses for Fairbanks rhyme test for subjects with normal hearing and subjects with sensorineural hearing loss. In L. L. Kopra (Project Director), Measurement of Speech-Sound Discrimination Loss, Terminal Progress Rpt., PHS Grant NB 04308 (1967).

LAWRENCE, D. L., and BYERS, V. W., Identification of voiceless fricatives by high frequency hearing impaired listeners. *J. Speech Hearing Res.*, **12**, 426-434 (1969).

OWENS, E., BENEDICT, M., and SCHUBERT, E. D., Consonant phonemic errors associated with pure-tone configurations and certain kinds of hearing impairment. *J. Speech Hearing Res.*, **11**, 656-667 (1972).

OYER, H. J., and DOUDNA, M., Structural analyses of word responses made by hard of hearing subjects on a discrimination test. *Arch. Otolaryng.*, **70**, 357-364 (1959).

PICKETT, J. M., MARTIN, E. S., JOHNSON, D., SMITH, S. B., DANIEL, Z., WILLIS, D., and OTIS, W., On patterns of speech feature reception by deaf listeners. In G. Fant (Ed.), *Proc. International Symposium on Speech Communication and Profound Deafness, Stockholm (1970)*. Washington: A. G. Bell Assoc. (1972).

ROSEN, J., Phoneme identification in sensorineural deafness. Unpublished Ph.D. thesis, Stanford Univ. (1962).

SIEGENTHALER, B. M., A study of the relationship between measured hearing loss and intelligibility of selected words. *J. Speech Hear. Dis.*, **14**, 111-118 (1949).

SIEGENTHALER, B. M., Perception of certain phonetic variables. *Speech Monogr.*, **21**, 39-45 (1954).

STUDDERT-KENNEDY, M., Speech perception. In N. J. Lass (Ed.), *Contemporary Issues in Experimental Phonetics*. New York: Academic Press (1976).

WANG, M. D., REED, C. M., and BILGER, R. C., A comparison of the effects of filtering and sensorineural hearing loss on patterns of consonant confusions. *J. Speech Hearing Res.*, **21**, 5-36 (1978).

WOLSTENHOLME, G. E. W., and KNIGHT J., *Sensorineural Hearing Loss*. London: J. and A. Churchill (1970).

Chapter II

PREVIOUS RESEARCH ON LINEAR AMPLIFICATION

A. PRELIMINARY REMARKS

It is obvious that amplification must be included in any signal process-ing scheme designed to improve speech perception for impaired listeners who have elevated absolute thresholds. It is equally obvious, however, on both theoretical and empirical grounds, that amplification by itself cannot restore normal or near normal speech perception for a large class of people with sensorineural impairments.

Among the theoretical arguments pointing out the inadequacy of amplification are the following. First, in many sensorineural impairments the auditory area is reduced because the elevation of absolute threshold is not accompanied by an equivalent elevation in the "saturation threshold" (for example, discomfort or pain threshold). Thus, independent of how the amplification is chosen (both with respect to overall gain and dependence of gain on frequency), it cannot restore the normal relationships between stimulus intensity and loudness within the auditory area. Second, as noted above, many sensorineural impairments are accompanied by loss of reso-lution and/or abnormal sound distortion within the auditory area. These phenomena are evident both in the subjective reports of impaired listen-ers and in the results of objective psychoacoustic tests. Finally, listeners with sensorineural impairments are typically required to listen to sounds at intensity levels that are sufficiently high to cause difficulties (distortion and spread of masking) even in normal listeners. Although the extent to which these problems can be alleviated by any form of signal processing is uncertain, it is obvious that amplification by itself is inadequate.

Empirical results on the effects of amplification on speech perception for listeners with sensorineural impairments are available in a wide vari-ety of sources, some of which are reviewed in detail below. In general, these results indicate that, although many listeners with sensorineural impairments benefit substantially from amplification, amplification by it-self does not provide them with normal (or near normal) speech under-standing abilities. Sometimes, the limitations of amplification are clearly evident even under ideal listening conditions. In other cases, the lim-

itations become evident only when interfering signals or reverberation are present. Among the results that demonstrate the inadequacy of amplification are the following. First, many listeners with sensorineural impairments who use hearing aids (most of which are merely amplification devices) are still unable to understand speech adequately. Although these aids are limited in the "quality" of amplification provided (for example, frequency response, distortion, and internal noise), and the procedures used to select an aid for an individual listener and to train him in the use of the aid are also limited, it seems unlikely that these limitations are the sole cause of the problem. Second, many listeners with sensorineural impairments who have participated in speech intelligibility tests (using relatively high-quality amplifying systems) in conjunction with audiological research or clinical examinations are unable to perform normally even when the overall gain of the system is chosen optimally. For example, many listeners with sensorineural impairments have intelligibility functions that are concave downward and have a peak value well below 100%. Third, in studies designed to determine the optimal frequency-gain characteristic to aid speech perception, a significant number of impaired listeners are unable to achieve normal speech perception even with the best characteristic studied. Although these empirical results do not prove that no linear amplification scheme could be designed to enable listeners with sensorineural impairments to perform normally, they are nevertheless consistent with this conclusion.

Given that amplification by itself is inadequate, it is nevertheless important to examine the results that are obtained with this form of processing. Any signal processing scheme that attempts to compensate for elevated absolute thresholds must include amplification. Furthermore, an understanding of the limitations of amplification may contribute to the design of improved schemes. Finally, more complex signal processing schemes will inevitably be evaluated by comparing their performance with that of schemes based simply on amplification.

Research on linear amplification for hearing impairments has concentrated in three areas: speech perception errors made with amplification, dependence of speech intelligibility on presentation level (or signal-to-noise ratio), and dependence of intelligibility on frequency-gain characteristic. With few exceptions, past studies fall into one, but only one, of these areas. In this review we focus our attention on the third area of research: the effect of varying the frequency-gain characteristic.

By the "frequency-gain characteristic of a linear amplification system" we refer to the acoustic transmission characteristic from sound source to eardrum referenced to the transmission characteristic that exists without the amplification system. Thus, to say that an amplification system has a flat or uniform frequency-gain characteristic means that the introduction of the amplification system does not alter the shape of the reference transmission characteristic; it does not mean that the system produces a transmission characteristic that is independent of frequency. Moreover,

since the acoustic transmission characteristic, even without the introduction of an amplifying system, depends on such factors as the angle of the source relative to the listener's head and the properties of the space in which the transmission takes place, the definition of frequency-gain characteristic must include a specification of the reference condition. In one such condition, often referred to as "orthotelephonic" (for example, Inglis, 1938; French and Steinberg, 1947; Richards, 1973), the listener faces the source at a distance of one meter in a free field.

Unfortunately, in many of the studies concerned with the effects of frequency-gain characteristic on speech intelligibility for impaired listeners, the frequency-gain characteristics have been specified inadequately. Thus, for example, when a characteristic is specified as flat, it often means that the frequency response of the amplification system is flat when the microphone is placed in a free field and the output of the earphones is measured by use of a coupler (rather than that the amplification system does not alter the frequency response that would be obtained without amplification). In order to ensure that the characteristic is *functionally* flat, it is necessary either to measure the change in the transmission characteristic under the actual conditions of use (for example, by obtaining aided and unaided field audiograms) or to take proper account of the differences between such measurements and the measurements employed (for example, the effects of head and pinna diffraction and concha and ear-canal resonances). Ideally, in discussing the various studies we would have attempted to correct for these factors; however, the studies do not generally contain sufficient information to permit such corrections. On the whole, the failure to make appropriate measurements or to take account of the relevant acoustic effects is most important for studies that employ insert earphones and electronics which amplify a broad range of frequencies, and least important for studies that employ circumaural earphones and electronics which amplify only a narrow band of frequencies. General information on relevant acoustic effects and the problem of system calibration is available in a wide variety of articles (for example, Dunn and Farnsworth, 1939; Romanow, 1942; Nichols et al., 1947; Shaw, 1966, 1973, 1974; Villchur, 1969; Zwiolocki, 1971; Burkhard and Sachs, 1975; Pascoe, 1975; Killion, 1976; Smaldino, Burkhard, and Chial, 1977; Burkhard and Sachs, 1977). Also, further material on the problem of adequately specifying the frequency-gain characteristic is presented in later sections of this chapter.

Another deficiency that appears in many of the studies is inadequate consideration of the problem of selecting and varying the levels of the speech materials used in comparative tests of different amplification systems. Except for internal amplifier noise and nonlinear distortion effects, the signal transformation produced by linear amplification is independent of signal level. However, when two different types of signal processing are compared (for example, different frequency-gain characteristics) it is essential that performance be described as a function of presentation level

so that comparisons can be made on the basis of maximum scores (over levels), maximum scores at comfortable levels, scores averaged over typical input levels, or operating range. Unfortunately, in many cases the input level has not been varied systematically or over an adequate range. Obviously, comparisons based on a single presentation level (specified either in SPL or SL) can be misleading because the performance-intensity characteristics for different forms of signal processing often have different shapes. Comparisons based on ranges of levels can also be misleading unless the levels leading to maximum scores for all of the systems being compared are included in the range of levels tested. In general, this level problem would appear to be more severe for studies that compare systems of widely different bandwidths at fixed objective levels than for studies that compare systems of comparable bandwidths at fixed subjective levels.

The effect of varying the frequency-gain characteristic on speech perception in impaired listeners has been studied by altering the characteristic of laboratory test equipment, by using a variety of commercial aids with different electrical characteristics, and by employing a variety of microphone-earphone-earmold configurations with different acoustic properties. Despite its long-recognized importance, however, the problem of determining the optimum characteristic for a given impaired listener remains unsolved. This failure is evident in the inconsistencies among the results of past research, in the lack of general agreement concerning the value of individual fitting for hearing aids or the manner in which such fitting should be accomplished, and in the existence of current research projects addressed to this problem. Among the factors which appear to have prevented the development of satisfactory techniques for determining this characteristic are: inadequate characterization of hearing impairments; inadequate materials and techniques for speech testing; inadequate consideration of relevant acoustic effects in specifying the functional (as opposed to nominal) characteristic of the amplifying system; inadequate consideration of interaction between the frequency-gain characteristic and other properties of the amplifying system, such as non-linear distortion; inadequate consideration of the effects of background interference and reverberation; and inadequate consideration of the effects of exposure time and training. Also, of course, the range of characteristics that have been tested is limited.

In the remainder of this chapter we review a wide variety of studies concerned with the effects of varying the frequency-gain characteristic on speech intelligibility for impaired listeners. In the 40 years spanned by these studies, significant changes have occurred in the testing equipment, techniques, and materials, in the characterization of hearing impairments, and in the population studied. Many of the early studies used articulation tests designed for highly trained listeners, described hearing loss solely in terms of air- and bone-conduction thresholds, and dealt with listeners suffering conductive, mixed, and sensorineural losses. In the course of time, articulation tests have given way to phonetically balanced word tests care-

fully screened for clinical populations, as well as a variety of special tests. Standard air conduction thresholds have been adopted and more careful techniques for measuring bone conduction thresholds have been developed. Also, threshold measurements are now often supplemented by a variety of other measurements in the attempt to characterize the loss more adequately. In addition, due to the medical advances in treating conductive impairments, current studies tend to focus more on listeners with impairments that are primarily sensorineural. Despite these changes, however, many of the early studies continue to be of interest.

For purposes of discussion, we have divided our review of previous research on the frequency-gain characteristic into five sections: (B) Studies of Filtering, (C) Studies of Selective Amplification, (D) Studies of Hearing Aids, (E) Studies Employing Children, and (F) Concluding Remarks. The material on adult listeners is separated from that on children because of the differences in the nature of the hearing losses studied in the two groups and differences in measurement problems. The studies of filtering and selective amplification are distinguished from the studies of hearing aids in that the former are concerned specifically with the effects of frequency-gain characteristic, whereas the latter are concerned with the effects of a wide variety of electroacoustic properties. In addition, most of the former studies were performed with relatively well-controlled, flexible, high-quality, laboratory equipment rather than commercial hearing aids. Finally, most of the studies of filtering differ from the studies of selective amplification in that they are concerned primarily with evaluating the contributions of different frequency bands to intelligibility rather than with comparing reasonable candidates for a useful frequency-gain characteristic. Accordingly, whereas in the section on filtering the useable range of frequencies is restricted by characteristics having relatively sharp slopes, in the section on selective amplification the slopes of the characteristics are relatively mild. Also, the studies of filtering, unlike those of selective amplification, often include material on normal listeners for comparison purposes.

B. STUDIES OF FILTERING

Studies of the effects of filtering on speech intelligibility have generally been performed in order to evaluate the contributions of different frequency bands to intelligibility and to determine the interactions among different bands.

Our discussion of these effects is divided into three subsections: (1) Highpass and Lowpass Filtering, (2) Bandpass Filtering, and (3) Multiple Disjoint Bands. In each subsection, we first consider results on normal listeners and then on impaired listeners. For both types of listeners, the studies have usually employed test stimuli consisting of monosyllables and have evaluated performance in terms of percent-correct identification. However, studies of normal listeners (except when they are used merely

as a control for the study of impaired listeners) have often differed from the studies of impaired listeners in that they have employed more precise stimulus control, better trained speakers, better trained listeners, and more systematic variation of both the filtering condition and the intensity level at which the material is presented. Many of the studies using normal listeners have been performed to evaluate transmission characteristics of communication systems and to explore the normal relationship between the acoustic and perceptual properties of speech, not merely as background or control for studies of impaired listeners.

The filtering studies considered, together with some of their relevant features, are listed in Table 1. Unless otherwise stated, the slopes of the filters used in these studies equal or exceed 24 dB/octave.

(1) Highpass and Lowpass Filtering

Research on the effects of highpass and lowpass filtering of monosyllables for normal listeners is exemplified by the studies of French and Steinberg (1947), Pollack (1948), and Miller and Nicely (1955).

French and Steinberg (1947) showed that when the unfiltered speech level is held fixed, performance decreases monotonically with bandwidth when either highpass or lowpass filtering is introduced in the transmission circuit. The crossover frequency, at which equal scores are achieved with highpass or lowpass filtering, was found to increase from 1700 to 1900 Hz as the presentation level increased. As Miller (1951) has pointed out, however, the 1900 Hz value is an average for male and female speakers: at high levels, the value for males is 1660 Hz and for females 2140 Hz. Within the framework of articulation theory (Collard, 1930; French and Steinberg, 1947; Beranek, 1947; Fletcher and Galt, 1950; Kryter, 1962a), the crossover frequency divides the spectrum into two bands that contribute equally to intelligibility. [Black (1959) obtained roughly consistent results using PB-50 words presented in both open response set (write-down) and closed response set (multiple-choice) test procedures.] French and Steinberg also found that for a fixed filter configuration, scores increase with signal level over a range of 60-70 dB, reaching a maximum at a gain of 0-10 dB relative to the level under the orthotelephonic condition, and then decreasing slightly at higher intensities.

Pollack (1948) reported that when such tests are carried out in a relatively intense (82 dB SPL) broadband noise background rather than in quiet, the results are modified in several important ways. For a fixed filter configuration, performance increases from chance to peak level over a range of only 25-35 dB. Also, the crossover frequency depends strongly on speech level, ranging from 800 Hz at low speech-to-noise ratios to 1620 Hz, close to the quiet value, at high ratios. Furthermore, scores in the noise background were not found to decrease monotonically as low-frequency speech components were removed by filtering: except at low

TABLE 1. Studies of filtering employing adult listeners.

Study	Date	Listeners[1]	Filters[2]	Q/N[3]	Voices[4]	Test Materials[5]
French and Steinberg	1947	8 N	HP, LP	Q	M, F	Nonsense CVC
Pollack	1948	9 N	HP, LP	N	M	PB-50[6]
Miller and Nicely	1955	5 N	HP, LP	N	F	Nonsense CV
Hirsh, Reynolds, and Joseph	1954	5 N	HP, LP	Q	M,F	Varied
Palva	1965	961 N	HP, LP, 1-2BP	Q	M	Finnish Trochees
Giolas and Epstein	1963	N	LP	Q	M	PB-50[6], W-22[7], Discourse
Giolas	1966	N	LP	Q	M	Sentences, Discourse
Speaks	1967	N	HP, LP	Q	?	SSI[8]
La Benz	1953, 1956	22 N, 78 I	HP, LP, BP	Q	M	PB-F[9], PB-K[10]
Parker	1953	10 I	HP	Q	M	PB-50[6]
Thomas and Pfannebecker	1974	9 I	HP	Q	M	PB-50[6]
Castle	1963	15 N	1BP	N	M	Nonsense CVC
Egan and Wiener	1946	6 N	1BP	Q	M	W-22[7]
Ambrose	1972	18 N, 36 I	1BP	Q	M	W-22[7]
Ambrose and Neal	1973	10 N, 10 I	1BP	Q	M	Nonsense CVC
Kryter	1962	3 N	1-3BP	Q	M	FRT[11]
Franklin	1969	36 N	2BP	Q	?	FRT[11]
Rosenthal, Lang, and Levitt	1975	4 N	2BP	Q	M	FRT[11]
Huizing and Taselaar	1961	? N	1-2BP	Q	?	Dutch PB Words
Linden	1964	50 I	1-2BP	Q	?	Swedish Spondees
Franklin	1975	6 I	1-2BP	Q	M	FRT[11]
Castle	1964	9 I	1-3BP	Q	M	PB-50[6]

[1] N = Normal Hearing, I = Impaired Hearing.
[2] HP = High Pass, LP = Low Pass, 1BP = Single Band Pass, 2BP = Two-Band Pass, etc.
[3] Q = Quiet, N = Noise Background.
[4] M = Male Voice, F = Female Voice.
[5] Consult individual studies for details of tests.
[6] Egan (1948).
[7] Hirsh et al. (1952).
[8] Speaks and Jerger (1965).
[9] Hudgins et al. (1947).
[10] Haskins (1949).
[11] Fairbanks (1958).

speech-to-noise ratios, better scores were obtained when frequencies below 350 Hz were removed before presentation.

Miller and Nicely (1955) studied the pattern of errors made in identifying filtered CV nonsense monosyllables. Under highpass filtering conditions, the observed patterns were found to be chaotic and unpredictable. Under lowpass conditions, however, the patterns were highly structured and indicated that the perception of the articulatory features is affected differentially. As progressively more high-frequency components are removed, the feature place is first affected, then duration and affrication; the features nasality and voicing can be perceived with good accuracy even when only low-frequency components are available. The variation in error pattern that was observed as the speech-to-noise ratio was decreased is similar to that for progressively more severe lowpass filtering and is consistent with the notion that a uniform spectrum noise masks high-frequency speech components more than low-frequency components.

The effect of filtering on different types of test materials has received comparatively little attention. In part, this appears to reflect the long-held belief (for example, Campbell, 1910) that correct reception of speech segments of syllabic size is sufficient (if not necessary) to assure correct reception of messages composed of larger speech segments. In this respect, the use of monosyllabic words represents a compromise between the use of nonsense syllables and the use of materials more like real speech (for example, sentences and connected discourse). Among the studies which have compared the effects of highpass and lowpass filtering on different types of test materials are those of Hirsh et al. (1954), Palva (1965), Giolas and Epstein (1963), Giolas (1966), and Speaks (1967). Hirsh et al. (1954) used word lists consisting of items having different syllabic structures. Recognition scores for polysyllabic and disyllabic words were found to be more resistant to highpass and lowpass filtering than were scores for monosyllables. Roughly the same crossover frequency was obtained for the three different types of test materials, however. Palva (1965) obtained somewhat variant results using materials consisting entirely of Finnish trochaic disyllables. For these materials, the crossover frequency was found to be 1100 Hz, significantly lower than the comparable value for monosyllables. Palva also found an unusually rapid drop in performance when components in the 500-600 Hz range were removed by lowpass filtering. He interpreted these results and others, discussed below, obtained for bandpass filtering, as indicating that performance on his test materials was more critically dependent on vowel discrimination than in comparable studies using English language materials.

Giolas and Epstein (1963) and Giolas (1966) compared the effects of lowpass filtering on words, sentences, and continuous discourse. Word scores were generally affected more by filtering when presented in isolation than when presented in sentence contexts. Also, scores for words in isolation were found to be poor predictors of comprehension scores on a test of the ability to relate the general content of discourse (Ulrich, 1957).

By contrast, scores for words in sentences were found to agree closely with comprehension scores for all but the most severe filtering conditions (540- and 780-Hz cutoffs). These results are often cited as evidence that performance on monosyllabic word lists is not a good indicator of performance on material more like real speech. This conclusion is not, by itself, surprising: the physical characteristic of a spoken word depends upon the context in which it is articulated (for example, Klatt and Stevens, 1973) and also word tests do not evaluate a listener's ability to utilize contextual cues. Nevertheless, the actual evidence provided by the results of these studies is far from conclusive. The relative predictive power of the word and sentence tests for performance on the comprehension test would have been altered if a more difficult comprehension test, or one more closely keyed to individual word reception, had been used. Also, the predictive power of the word tests was evaluated using an extremely limited concept of how the results of one test can be used to predict the results of another test. For example, the possibility that a valid but non-linear predictive relation between the word scores and the comprehension scores existed, was not examined.

Speaks (1967) evaluated the effects of both lowpass and highpass filtering on the identification of synthetic sentences. In the test procedure used, the materials consisted of a single set of 10 seven-word messages constructed to be statistical approximations to English sentences. Listeners were informed of both the complete message set and of the correctness of each identification response. The results are similar to those obtained for monosyllable identification in that at a given presentation level, performance decreases as the severity of filtering increases. The results are dissimilar, however, in that for each filtering condition performance is error free at sufficiently high input levels (for example, at 65 dB SPL for the 125-Hz lowpass condition and at 60 dB SPL for the 7000-Hz highpass condition). Also, the range of intensities from near-chance to near-perfect performance is relatively small (10-20 dB) for all filtering conditions. This interaction of level and bandwidth was attributed to the greater redundancy of sentence-like materials. Because perfect identification performance could be achieved at sufficiently high intensities for each filtering condition, it was possible to determine the crossover frequencies only at very low signal levels (15-20 dB SPL) for these materials. Speaks interpreted the low value thus obtained (725 Hz) as assigning considerably more importance to the low-frequency components of speech than is indicated by studies based on monosyllables.

This study has often been cited as evidence that tests based on monosyllables or words do not provide an accurate indication of performance on materials more like real speech. In interpreting the results of the synthetic sentence identification procedure, however, it is important to note that the set of possible messages was small, closed, and known in advance, unlike most real speech. Consequently, nearly perfect message identification could be achieved under conditions in which it is unlikely

that a single word in the message would have been identified correctly. Clearly this effect depends critically on the use of a highly restricted message set, as well as on the use of sentence-like materials. It is very unlikely that results so disparate from the findings for monosyllables would have been obtained with much larger message sets (for example, sets containing 10,000 sentences).

As stated above, studies of the effects of filtering on impaired listeners have generally been less controlled and systematic than studies of normal listeners. Among the more interesting studies of highpass and lowpass filtering involving impaired listeners are those of La Benz (1953, 1956), Parker (1953), and Thomas and Pfannebecker (1974).

La Benz (1953, 1956) tested three groups of impaired listeners (conductive, sensorineural, and mixed) and a control group of normal listeners. On the average, those with sensorineural impairments had losses that increased from roughly 20 dB at 250 Hz to 60 dB at 5000 Hz. Materials were presented 30 dB above each listener's SRT (that is, 30 dB above the threshold for unfiltered spondees). When filtering (slopes of 30 dB/octave) was used, the gain was increased to maintain roughly constant delivered speech power. The scores obtained for the unfiltered condition by the listeners with sensorineural impairments were lower on the average than those obtained by the control group (73% compared to 95%). However, the differences between groups decreased as progressively more severe lowpass filtering was introduced. For the 500-Hz lowpass condition, virtually identical scores were obtained by all groups (roughly 30%). In contrast, highpass filtering increased the differences between groups. Under the most severe highpass filtering tested (2000-Hz cutoff), the group with normal hearing obtained nearly the same score as in the unfiltered case (92%), but the average score for the sensorineurals was only about 60% of its unfiltered value. For the sensorineurals, La Benz found a crossover frequency of roughly 1000 Hz, considerably below the value for listeners with normal hearing.

In general, La Benz's results support the expectation that listeners with sensorineural impairments who have falling audiograms with mild losses at the lower frequencies can make normal (and only normal) use of cues conveyed by low-frequency speech components but not of cues conveyed by high-frequency components. One result that is controversial in the light of other studies, however, is that for all groups tested, the average performance for each group decreased consistently as progressively more highpass filtering was introduced. Many other studies suggest intelligibility can be increased by attenuating the lower frequencies. In La Benz's study, for the 250- and 500-Hz highpass conditions, 10 of the 22 sensorineurals showed decreased scores, 11 showed no significant changes, and only one showed improved performance relative to the wideband condition. Those whose scores decreased generally had greater losses above 1000 Hz (averaging 10 dB worse in the 3000-6000 Hz range) than those whose scores were roughly unchanged. Both of these subgroups,

however, had very similar losses below 1000 Hz. The one listener whose score improved significantly had no hearing loss up to 2000 Hz, but roughly 60 dB loss at 3000 Hz and above. The results obtained by La Benz on bandpass filtering are considered in the next section.

Parker (1953), in an exploration of the hypothesis that the limited speech intelligibility evidenced by listeners with sensorineural impairments is the result of strong speech sounds causing an abnormal threshold shift that persists through the following sounds, studied the effect of a variety of speech transformations, including highpass filtering at 670 Hz (50 dB down at 470 Hz). He tested first without filtering over the intensity range SRT + 6 dB to SRT + 36 dB. He then retested with filtering at the unfiltered input level that had yielded the best score for each subject. The group average score with filtering was 5 points higher than in the best unfiltered condition (62% compared to 57%). Four of the subjects showed dramatic improvements with filtering (on the average, 47 to 68%), four showed only minor improvements (69 to 77%), and two showed substantial decreases in score (56 to 20%). On the average, all three groups had equivalent losses at 250 Hz, those who showed subtantial improvements had losses which increased roughly 12 dB per octave above 500 Hz, those who showed minor improvements had relatively flat audiograms, and those who showed decreases had much greater losses at frequencies of 500 Hz and above than the other groups (for example, a loss of roughly 90 dB at 2000 Hz). Although Parker's results for the two subjects with very large losses above 500 Hz are to be expected (as Parker himself points out) and are consistent with the results of La Benz, we see no obvious explanation for the differences between the results for Parker's other subjects and the results of La Benz.

Thomas and Pfannebecker (1974), in a study designed to explore the possible benefits of reducing the masking of higher-frequency components (particularly the second formant) by energy at the first formant, explored the dependence of speech intelligibility on the attenuation slope of a highpass filter. Based in part on the results of Thomas and Ohley (1972) on the intelligibility of speech in noise for normal listeners, the cutoff frequency and slopes were chosen to be 1600 Hz and 12, 18, and 24 dB/octave. The tests were applied to 9 subjects with an average hearing loss of roughly 25 dB at 250 Hz and 70 dB at 4000 Hz. (Etiologies of the losses included noise exposure, presbycusis, and otosclerosis.) Presentation levels for the unfiltered material were SRT + 20, 30, and 40 dB. For the filtered conditions corresponding to each of these unfiltered levels, the amplitude of the filtered speech was increased so that the delivered speech power was the same for the filtered and unfiltered conditions. In almost all cases (that is, independent of the subject, the slope, and the level), the score for the filtered speech was significantly better than for the unfiltered speech. For example, at SRT + 30 dB, the improvement in score afforded by the filtered speech (averaged over subjects and slopes) was roughly 15 points (52% to 67%). The dependence of performance on

the slope of the filter was not as strong as on the presence of the filter and varied considerably as a function of both the listener and the level. In interpreting these results, which are more consistent with the results of Parker than La Benz, it is important to note that only a limited range of levels was tested and that seven of the nine subjects achieved maximum scores for the unfiltered material at the highest level tested. Thus, it is unclear whether the improvement afforded by the filtering would have been sustained if higher levels had been tested.

In general, in considering the above results, it should be noted not only that many of the results are inconsistent, but that all of the above tests on impaired listeners were conducted in quiet. It is obvious that the effects of lowpass or highpass filtering can be substantially altered by the introduction of background noise. Thus, for example, if low-frequency room noise is added to the speech material, highpass filtering is likely to prove more beneficial. Further results on lowpass and highpass filtering are included in later sections.

(2) Bandpass Filtering

Illustrative results on the effects of bandpass filtering for normal listeners have been obtained by Castle (1963), Egan and Wiener (1946), and Palva (1965); and for both normal and impaired listeners by La Benz (1953, 1956), Ambrose (1972), and Ambrose and Neal (1973). (Since Castle's study of impaired listeners includes the use of multiple disjoint bands, it is considered in the next section.)

Castle (1963), using filters with relatively steep slopes, obtained word recognition scores in excess of 90% with normal listeners for a bandwidth of 1440 Hz centered on 1080, 1680, or 1920 Hz at comfortable levels. La Benz's (1953, 1956) results on normals under bandpass conditions are comparable, allowing for differences in the definition of bandwidth associated with filter slope. Data obtained by Egan and Wiener (1946) on bandpass filtered speech (using extremely sharp filters) in the presence of broadband noise suggest, however, that these results may apply only at very high signal-to-noise ratios. For example, in the presence of an 84 dB SPL uniform spectrum noise, a speech level adequate to achieve a score of 90% for unfiltered speech (a level of 35 dB above the orthotelephonic reference) results in a score of roughly 45% when the frequency range of the speech signal is restricted to 550-1500 Hz. The results of Egan and Wiener (1946) have been interpreted (Licklider and Miller, 1951) to suggest that, within limits, bandwidth and presentation level are tradable. However, results for very narrowband speech presented at very high levels are not available.

Palva (1965) conducted an extensive series of bandpass-filtering tests using materials consisting of Finnish trochaic disyllables. He employed filters with center frequencies ranging from 240 to 5000 Hz and

bandwidths ranging from ¼ to 1-½ octaves. For each filter condition, materials were presented at 50 dB SL. He found consistent, with the results of Castle (1963), that, for a fixed center frequency, performance improved as bandwidth increased. For fixed bandwidth, however, he found that performance was best for center frequencies near 1500 Hz, second best for center frequencies near 600 Hz, poor for center frequencies near 800 Hz, and very poor for frequencies at the edges of the frequency range tested. Palva interpreted these results as indicating that relatively good recognition scores can be achieved when the filter passband coincides with either the first or second vowel formant regions.

La Benz (1953, 1956) obtained results on both impaired and normal listeners using bandpasses of 200-500, 250-750, 500-1000, 750-1500, 2000, and 2000-3000 Hz (the filter slopes, levels, and subjects were the same as in the highpass-lowpass study cited previously). While the scores obtained by normals increased steadily from roughly 45 to 85% as the center frequency of the band increased, the scores obtained by the sensorineurals remained relatively constant and never exceeded 45%. However, in considering these results (as well as the results of other filtering studies in which only one level was tested for each subject and this level was chosen without regard for the shape of the audiogram), it should be noted that the tests for different subjects and different bands were performed at different sensation levels. For example, when the 250-750 Hz band was used, the average presentation level was 25 dB SL for the normals and 35 dB SL for the sensorineurals; and the average scores were 44 and 34% respectively. When the 1500-2000 Hz band was used, the corresponding levels were 36 and 21 dB SL and the scores were 75 and 37% respectively. Thus, at least a portion of the difference between the normals and the sensorineurals with respect to the dependence of intelligibility on the frequency band can probably be ascribed to differences in sensation level.

Ambrose (1972), in a study concerned with the effects of harmonic distortion, studied both normal and impaired listeners (conductives and sensorineurals) using bandpasses of 600-1200, 1200-2400, and 300-4800 Hz, filter slopes of 90 dB/octave, and presentation levels of SRT + 40 dB (maintaining constant delivered power for all conditions). Independent of the amount of harmonic distortion, the sensorineurals (who had falling audiograms) obtained roughly the same low scores as the normals for the lowband case, but much worse scores than the normals for the highband and broadband cases. Also, the scores obtained by the sensorineurals for the wideband case were vastly superior to the scores they obtained for the two narrowband cases. The results of Ambrose, like those of La Benz, support the notion that sensorineurals with falling audiograms can make normal use of low-frequency cues, but not high-frequency cues. However, it does not follow from these results, as implied by Ambrose, that low-frequency emphasis will increase intelligibility more than mirroring the audiogram (that is, using a frequency-gain characteristic that makes the

aided absolute threshold curve for the impaired listener identical to the unaided absolute threshold curve for normal listeners).

Ambrose and Neal (1973) studied normals and sensorineurals using bandpasses of 200-1500 (L), 500-2000 (M) and 1000-3000 (H) Hz, filter slopes of 24 dB/octave, and presentation levels of SRT + 40 dB (with gain adjusted to maintain constant delivered power for all filtering conditions). The normal listeners obtained high scores for all filtering conditions (the average score for each bandpass exceeded 90%), although performance was best for the H filter. The impaired listeners, whose average audiogram showed a 45 dB loss at 250 Hz and an 80 dB loss at 4000 Hz, produced scores ranging from 2 to 94%, with an average best score of 75%. The seven impaired who did best in the H condition had generally greater than average losses at low frequencies and less than average losses at high frequencies. By contrast, the eight who did best or second best in the M condition had smaller losses than average in the mid-frequency and high-frequency regions, and the six who did best or second best in the L condition had smaller losses than average in the low-frequency region and larger than average losses in the high-frequency region. Thus, differences in scores for the different filter conditions again appear to be, at least in part, a reflection of differences in the sensation levels at which the material was presented.

(3) Multiple Disjoint Bands

Illustrative results for normal listeners on the effects of filtering by means of multiple disjoint bands have been obtained by Kryter (1962b), Franklin (1969), Rosenthal, Lang, and Levitt (1975), Huizing and Taselaar (1961), Linden (1964), and Palva (1965). Many of these studies have included an evaluation of the effects of augmenting a high-frequency band with a low-frequency band and of the differences between presenting the low-frequency band to the same (ipsilateral) ear or to the opposite (contralateral) ear.

Kryter (1962b), in a study concerned with validation of the Articulation Index for normal listeners and employing filter slopes of 60 dB/octave and nonsense syllables as test material, found that (in quiet) adding the low-pass band 0-600 Hz to any of the bands 1200-2400 Hz, 1200-1700 Hz, and 1700-2400 Hz resulted in substantial increases in performance (62-85%, 30-70%, and 45-75%, respectively). However, essentially no further improvement was obtained by also adding the band 4800-9600 Hz.

Franklin (1969), who was concerned with the possible benefits of providing low-frequency information to listeners with little or no residual hearing above 500 Hz, studied the effect on consonant discrimination (Fairbanks Rhyme Test) by normal adults of adding a one-octave wide low-frequency band (240-480 Hz) to a one-octave wide high-frequency band (1020-2040 Hz) using filter slopes of 60 dB/octave and a presentation level of 0 dB SL for the high passband. The low passband was presented

at 0, 20, or 40 dB SL either ipsilaterally or contralaterally. Except for the 40 dB level, the results for the ipsilateral case were approximately the same as for the contralateral case and indicated that the addition of the low passband was helpful: when only the high passband was presented, the score was 41% (21 points above chance performance); when the low passband was added at 0 dB SL, the score was 55%; when it was added at 20 dB SL, the score was 62%. When the low passband was presented at 40 dB SL, masking occurred and the score dropped to 54% for the contralateral case and to 38% for the ipsilateral case (leading Franklin to suggest that the best aid for certain impaired listeners might be achieved by splitting the frequency spectrum between the two ears).

Rosenthal, Lang, and Levitt (1975), in a further study of the same type using normal adults and the Fairbanks Rhyme Test, examined the effects of adding a one-octave wide low passband (55-110, 110-220, or 220-440 Hz) either ipsilaterally or contralaterally to a one-octave wide high passband (1020-2040 Hz) using filter slopes of roughly 70 dB/octave. The high passband was presented at a constant level corresponding to unfiltered speech at roughly 80 dB SPL (which produced a score of roughly 34% when presented alone), and the low passband was presented at five levels from 0 to 40 dB re normal speech level (corresponding to unfiltered speech at 80-120 dB SPL). Averaged over the ipsilateral-contralateral distinction, the results indicate that the addition of the low passband was always helpful, the amount of improvement increased with the center frequency of the low passband, the improvement in score per cycle of bandwidth was greatest by a slight amount for the lowest passband tested, and the optimum level for the low passband was in the region 20-30 dB re normal speech level. The maximum improvement obtained was 34 to 64% and occurred when the 220-440 Hz band was presented at a level of 20 dB. Only at the highest level tested (corresponding to unfiltered speech at 120 dB SPL) did the contralateral presentation lead to a consistently higher score than the ipsilateral presentation. Although the results of this study are generally similar to those of Franklin with respect to the benefits of adding a low-frequency band, the detailed results of the two studies appear inconsistent with respect to the level parameter and the occurrence of masking.

Additional studies employing multiple disjoint bands have been performed by Huizing and Taselaar (1961), Linden (1964), and Palva (1965). A wide spectrum of results has been obtained both with respect to performance level and to the relative advantage of contralateral over ipsilateral presentation. For example, Huizing and Taselaar using a 140-280 Hz low-frequency band presented at 75 dB SPL found that a contralateral presentation permitted a 14 dB reduction in the presentation level of the high-frequency band (1128-2256 Hz) for the same performance. Linden and Palva found no contralateral advantage when the bands were presented at the relative levels characteristic of normal speech or at the same sensation level. However in each case the low-frequency band tested was

relatively high in frequency (560-715 Hz and 480-720 Hz) and the high-frequency band was relatively narrow (1800-2220 Hz and 1800-2400 Hz). In general, the results on multiple bandpass filtering indicate that supplementary low-frequency signals can improve word recognition, but that at high levels the improvement is counteracted by spread of masking effects. The detailed picture, however, is far from clear.

Results for impaired listeners on the effects of multiple disjoint bands have been obtained by Franklin (1975) and Castle (1964). Franklin, in an extension of her earlier work on normals, studied the effect of adding a low passband to a high passband on consonant discrimination by six impaired listeners with moderate-to-severe binaural sensorineural losses. The average loss for these listeners increased from roughly 50 dB at 250 Hz to more than 90 dB at 4000 Hz (for both the better ear and the worse ear). The high passband (1020-2040 Hz) was presented to the better ear at 10 dB SL and the low passband (220-440 Hz) was presented either ipsilaterally or contralaterally at 10, 30, or 50 dB SL. On the average, the high passband alone produced a score of 36% (chance = 20%). The addition of the low passband had essentially no effect when presented ipsilaterally, but led to significant improvement when presented contralaterally (increasing the score to slightly over 50%). Moreover, these results were roughly independent of the level at which the low passband was presented. These results differ from Franklin's results on normals in that the contralateral stimulation was consistently better than the ipsilateral stimulation, independent of the impaired listener and the level of the lowpass speech.

Castle (1964) explored a wide variety of multiple-disjoint-band filter combinations on nine listeners with sensorineural impairments using a presentation level of 35-40 dB above SRT and filter slopes of 30 dB/octave and greater. The average audiogram for the group showed a loss of roughly 20 dB at 250 Hz and 70 dB at 4000 Hz. Aside from the unfiltered condition, he studied filtering combinations that amplified most in regions of greatest sensitivity (matching), amplified most in regions of least sensitivity (mirroring), and provided selective narrowband amplification. Although six of the listeners obtained higher scores for one or another of the filtering conditions than for the unfiltered condition, only two of these listeners (those with trough-shaped audiograms) showed improvements that exceeded the large test-retest variations that occurred. Furthermore, the results for these two listeners were very curious (one achieved improvement for all filter combinations tested, including 1000-2500 and 60-500 plus 8000-20,000; the other showed a drop in score of 22% in going from 60-2150 Hz to 60-2280). Aside from noting that one cannot determine the best filter combination on the basis of the audiogram, Castle concludes from his results that some individuals are likely to be aided the most by restricting the amplification to one or more narrow bandwidths (for example, 1440-1680 Hz). In considering Castle's conclusions, however, it should be noted that the picture presented by his actual data is

extremely unclear; many of the differences among the scores for the different filter combinations on which he bases his conclusions are comparable to the test-retest variations that occurred.

C. STUDIES OF SELECTIVE AMPLIFICATION

Research on selective amplification differs from most research on filtering in that it is more directly concerned with the task of determining the optimum frequency-gain characteristic for a given impaired listener or group of such listeners. Accordingly, much less use is made of normals for comparison purposes and the slopes associated with the characteristics are usually less severe. Our discussion of this research is divided into two subsections: (1) Principles and (2) Experiments. In the first, we review some of the ideas and models that have been suggested to guide the choice of characteristic. In the second, we review results of studies that have compared a variety of characteristics experimentally.

(1) Principles

The principles that have been proposed to guide the choice of frequency-gain characteristic vary widely with respect to the properties of the listener's impairment and acoustic environment that are considered, the relative weights given to speech and nonspeech stimuli, and the amount of theoretical modelling involved. Not surprisingly, the characteristics implied by these principles for a given impaired listener vary widely with the principle. In the following paragraphs, we review a representative sample of these principles.

Among the ideas that have been proposed to guide the choice of frequency-gain characteristic (some of which are clearly contradictory) are the following. First, the gain should increase in frequency regions where the hearing loss increases so that the impaired listener's aided sensitivity approximates that of the normal. In the extreme version of this idea, the characteristic is chosen to be the mirror image of the audiogram so that the aided listener's absolute threshold curve is restored to normal. In a less extreme version, that takes account of the fact that most listening does not take place at absolute threshold and that the listener's dynamic range is often seriously reduced in frequency regions where the hearing loss is great, the characteristic is derived from an equal-loudness contour substantially above threshold (for example, Watson and Knudson, 1940). For cases with sloping audiograms such a characteristic has less high-frequency emphasis than one that mirrors the audiogram because equal-loudness contours tend to be flatter than absolute threshold curves. Second, the gain characteristic yielding best speech reception is relatively independent of the hearing loss (for example, Davis et al., 1947). In a special version of this idea, the gain is chosen to be independent of frequency so that the amplification system does not introduce external linear

distortions which compound the effects of the impairment (for example, Pascoe, Niemoeller, and Miller, 1974; Miller, 1974). Third, the gain should decrease in frequency regions where the loss is large (and the dynamic range is small) in order to reduce the effects of internally generated distortions associated with stimulation of these regions. For example, Pimonow (1963) suggests that if the loss in a given frequency region exceeds roughly 60 dB, this region should be attenuated relative to those of smaller loss, because amplifying this region will deteriorate overall intelligibility. Similarly, Huizing and Reyntjes (1952) suggest that even in cases of relatively uniform hearing loss, the characteristic should emphasize regions in which the "subjective distortion" is low and suppress regions of high internal distortion.

Radley et al. (1947) attempted to compute the characteristic that would yield the highest articulation score subject to a constraint on total delivered power (to prevent discomfort of the listener or overloading of the amplification system). According to their results, the optimum characteristic in quiet, G(f), is given by

$$G(f) = K \cdot A(f) \ [L(f)]^{(\alpha-1)/\alpha}$$

or, in dB,

$$20 \ \text{Log} \ G(f) = 20 \ \text{Log} \ K + 20 \ \text{Log} \ A(f) + [20 \ (\alpha-1)/\alpha] \ \text{Log} \ L(f),$$

where L(f) is the hearing loss at frequency f, α is a constant derived from an approximation to the slope of the band-articulation function, K is a constant derived from the constraint on total power, and A(f) is derived from the normal absolute threshold curve, the shape of the speech spectrum, and the relative importance of speech components at frequency f. When background noise is present, the optimum characteristic G(f) is modified by the inclusion of the multiplicative factor $[W(f)]^{1/\alpha}$, where W(f) is the speech-to-noise ratio at frequency f. In applying this theory, Radley et al. assumed that both A(f) and α are the same as for normals. According to their results, A(f) increases at roughly 6-8 dB/octave (for both male and female speakers) and $\alpha \approx$ ¾. Since $\alpha < 1$, the function G(f) provides less high-frequency emphasis for cases of falling audiograms than for flat audiograms. However, since α is close to unity, the function G(f) is dominated by A(f) and depends only weakly on the loss L(f). Also, the data used to determine α could be equally well fit by values of α slightly greater than unity. Finally, since the speech-to-noise ratio in many situations decreases at both low and high frequencies, the effect of background noise on the optimum characteristic is generally to provide more relative gain in the middle frequency region.

Fletcher (1952) analyzed the results of the study by Davis et al. (1947) in terms of articulation theory and concluded that the gain characteristic yielding best scores is given by

$$G(f) = L_C(f) \cdot [L_N(f)]^r,$$

where $L_C(f)$ is the conductive loss at frequency f, $L_N(f)$ is the nerve loss, and r is a constant in the range $0.2 < r < 0.4$. According to this formula, gain increases in regions of greater loss but not as rapidly as elevation of threshold when the loss is sensorineural. Roughly speaking, this gain characteristic is similar to that based on an equal-loudness contour well above threshold. Although it does not provide a test of the shape of the gain characteristic, a recent study by Martin (1973) indicates that the range of the exponent r suggested by Fletcher is consistent with the actual gain settings used by sensorineurals in conversational situations. For body-worn aids the average exponent was 0.4; for head-worn aids, 0.2. Roughly similar results, recently summarized by Martin and Grover (1976), have been reported by Brooks (1973), Byrne and Fifield (1974), and Martin et al. (1976).

De Vos (1968) proposed that the frequency-gain characteristic should be chosen such that the relationship between the speech area and the most comfortable equal loudness contour is restored to normal. The acoustic gain at frequency f is determined by the hearing loss $L(f)$, the sensation level of speech for a normal listener $S(f)$, and the fractional reduction in dynamic range $R(f)$, according to the formula

$$G(f) = L(f) \cdot [S(f)]^{-R(f)}.$$

According to this relation, the gain mirrors the audiogram only when the impairment is not accompanied by a reduction in dynamic range [that is, when $R(f) = 0$]. In cases of recruitment, $R(f) > 0$ and the indicated gain is less than that required to mirror the audiogram. Although no formal tests of this relation were provided, the method was claimed to be useful and reliable in "thousands" of fittings of persons who developed impairments at advanced ages. More recently, Barfod (1972), on the basis of intelligibility tests administered to 10 impaired listeners, has suggested a very similar basis for choosing the frequency-gain characteristic. Specifically, he indicated that the characteristic should transform the 10% cumulative levels of speech to have the same loudness as for normal listeners.

Victoreen (1973a) has suggested that the gain characteristic be chosen so that, to the extent possible, the average amplified speech spectrum for the impaired listener has roughly the same relation to the threshold of audibility and the threshold of discomfort as unprocessed speech presented at a comfortable level has for normals. More specifically, he recommends that the gain characteristic be chosen such that the speech spectrum is delivered at a level between one-third and one-half the distance from the threshold of audibility to the threshold of discomfort. This characteristic can be expressed as

$$G(f) = L(f) \cdot \left[\frac{M(f)}{L(f)}\right]^b,$$

where M(f) is the elevation of the discomfort threshold for the impaired listener and b lies in the range $0.3 < b < 0.5$. Under the assumption that the loss is sensorineural and that the discomfort threshold is the same as for normals, this formula is identical to Fletcher's with the constant $1 - b$ replacing the constant r. Note, however, that whereas the specified range for $1 - b$ is $0.5 < 1 - b < 0.7$, the specified range for r is $0.2 < r < 0.4$. Thus, Fletcher's recommended characteristic depends less strongly on the loss. Note also that unlike Fletcher's formula, the above formula requires a reduction in gain when the discomfort level is lower than normal.

Wallenfels (1967), who also stresses the importance of preventing sounds from becoming too loud (if an aid produces uncomfortably loud sounds it will simply not be worn), suggests that the spectrum of speech sounds be amplified to fall halfway between the threshold of audibility and the threshold of discomfort for frequencies above 1000 Hz. Below 1000 Hz, he suggests a number of alternatives based on the overall character of the loss, but in general advocates progressive attenuation of the low frequencies at a rate of 8-10 dB per octave or more.

Byrne and Tonnisson (1976) have argued that the frequency-gain characteristic should be adjusted so that "all frequency components of speech will be presented with approximately equal loudness so as to maximize the amount of signal that can be received through a hearing aid set at a comfortable listening level." This characteristic, according to the authors, is given by

$$G(f) = H(f) \, [L(f)]^\alpha,$$

where H(f) is a fixed function (based on normal hearing) chosen to adjust the loudness of ⅓ octave bands of speech to 60 phons, L(f) is the loss, and $\alpha = 0.46$. The function H(f) rises 18 dB from 250 to 1000 Hz, is roughly flat from 1 to 2 kHz, and falls roughly 6 dB from 2 to 4 kHz. The constant α is chosen to reflect the fact (noted above) that the actual gain settings used by impaired listeners are generally substantially lower than those required to compensate for the loss in sensitivity. Although Byrne and Tonnisson did not test this characteristic experimentally, they claim that, except for the inclusion of adjustments [H(f)] to the speech spectrum, their procedure is relatively consistent with that suggested by Watson and Knudsen.

Huizing, Kruisinga, and Taselaar (1960) have advocated selection of the frequency-gain characteristic on the basis of "triplet audiometry." In this technique, speech reception is measured for three filtering conditions: low-pass with a 900-Hz cutoff, octave-bandpass from 900-1800 Hz, and high-pass with an 1800-Hz cutoff. The frequency-gain characteristic

selected for a given listener depends upon the pure tone audiogram, the "triplet-audiometry" discrimination scores, and the temporal history of the loss. If reception is good for all three bands, amplification is selected to compensate for the hearing loss as indicated by the pure-tone audiogram in order to prevent possible deterioration of "cerebral patterns" associated with atrophy of sensory mechanism or brain centers. If reception is poor for a given band (score of 20% or less) and there is better discrimination in other bands, then the poor band is not amplified. According to this theory, however, neither the audiogram nor the test scores determine the frequency-gain characteristic uniquely. The choice of characteristic is also influenced by the age of the listener and the suddenness of onset of the loss (which are assumed to affect the likelihood that "cerebral patterns" can be excited by amplified speech sounds).

Guberina (1972) and Asp (1975) have advocated the use of the "optimal field of hearing" as a guide in determining the frequency-gain characteristic. According to this theory, which bears certain similarities to triplet-audiometry, the frequency regions to be amplified include those to which the impaired listener is most sensitive plus any additional regions needed to understand speech. The rationale for this choice is that impaired listeners depend on acoustic cues not utilized by normal listeners (Rhodes, 1966; Lawrence and Byers, 1969) and can learn to discriminate speech by using existing hearing rather than by resurrecting hearing that has not been used. The "optimal field of hearing" is determined by rather unconventional techniques: the detection threshold is measured using stimuli consisting of nonsense monosyllables filtered in octave-wide bands, and speech discrimination is measured using monosyllable recognition tests in which the stimuli are graded according to tonality.

Although many of the assumptions that form the basis of the suggestions advanced by Guberina, Asp, and Huizing et al. have not been formally tested, and the procedures advocated depend heavily on clinical skills and judgements, these efforts are among the few that include tests with filtered speech in the determination of the frequency-gain characteristic.

Recently Lawrence and Blackledge (1977) and Levitt et al. (1978) have begun to develop protocols for selecting the frequency-gain characteristic from among those provided by master hearing aids on the basis of adaptive sequences of speech reception tests.

(2) Experiments

Empirical studies of selective amplification, like the principles discussed above, are difficult to summarize and often contradictory. Our discussion of these studies, which are listed in Table 2, is organized historically.

Watson and Knudson (1940) measured syllable articulation in 16 impaired ears (9 conductives and 7 sensorineurals with a wide range of losses) using a variety of frequency-gain characteristics, including one de-

TABLE 2. Studies of selective amplification employing adult listeners.

Study	Date	Listeners[1]	Q/N[2]	Voices[3]	Test Materials[4]
Watson and Knudson	1940	16 I	Q	?	Monosyllabic Words
Davis et al.	1947	25 I	Q, N	M	PB-50[5]
Hudgins et al.	1948	6 I	Q, N	M	PB-50[5]
Sheets and Hedgecock	1949	105 I	Q	M	PB-50[5]
Radley et al.	1947	225 I	Q	M, F	Monosyllabic Words
Knight	1967	81 I	Q	M	Monosyllabic Words
Reddell and Calvert	1966	24 I	Q, N	M	W-22[6]
Katakura et al.	1968	43 I	Q	?	Syllables
Thompson and Lassman	1969	30 I	Q, N	?	W-22[6], SSC
Kemker	1972	24 I	Q	?	NU Test 20-N
Barfod	1972	15 I	N	M	Nonsense CVC
Johansen and Simonsen	1973	10 I	Q	?	"Current Speech"
Pascoe et al.	1974	1 I	Q, N	F	W-22[6]
Pascoe	1975	8 I	Q, N	M, F	W-22[6], H-F Lists
Skinner	1976	6 I	Q	M, F	H-F Lists

[1] I = Impaired Hearing.
[2] Q = Quiet, N = Noise Background.
[3] M = Male Voice, F = Female Voice.
[4] Consult individual studies for test details.
[5] Egan (1948).
[6] Hirsh et al. (1952).

rived from the most comfortable equal-loudness contour and one that was nominally flat. The former characteristic, which produced the best test results, can be expressed by

$$G(f) = K \cdot L(f) \cdot D(f) = K \cdot MC_I(f)/T_N(f),$$

where K is a constant, L(f) is the loss, D(f) is the most comfortable sensation level for the impaired listener, $MC_I(f)$ is the most comfortable intensity level for the impaired listener, and $T_N(f)$ is the normal absolute threshold level. For almost all the ears tested, but particularly for the sensorineurals, this characteristic was superior to the flat characteristic, producing an average score of 77% compared to 66%. [Unfortunately, the characteristic $G(f) = K \ MC_I(f)/MC_N(f)$, where $MC_N(f)$ is the most comfortable intensity level for normals, was not tested.] In addition to concluding that the above characteristic is superior to the flat characteristic, Watson and Knudson stress the importance of distinguishing between objective test results and subjective judgements of impaired listeners in determining the best characteristic. As illustrated by three sensorineurals with falling audiograms, listeners are likely to prefer characteristics that accentuate the frequencies that they can hear best without amplification. Also, Watson and Knudson demonstrate that objective test results may depend strongly on exposure time and training. For example, one listener with a large sensorineural loss produced performance curves (test score versus time) for selective and flat amplification that increase from roughly 10 to 40% during a three-month period and that cross each other.

Davis et al. (1947), in the well-known Harvard "Master Aid" study, compared a variety of frequency-gain characteristics using PDR-10 earphones and a basic amplification system that was flat within ±5 dB from 100 to 7000 Hz, as determined with a 6-cc coupler. The principal characteristics studied can be described by

$$G(f) = K \cdot f^{2n}$$

where $n = -2, -1, 0, 1,$ or 2 (corresponding, respectively, to attenuation of 12 and 6 dB per octave, uniform or flat gain, and high-frequency emphasis of 6 and 12 dB per octave). The speech tests were conducted over an amplitude range of 40-70 dB to determine the maximum score and operating range (that is, the range over which the score exceeded 50%) for each characteristic and listener. In most of the tests, the listeners were protected from overly intense sounds by symmetrical peak-clipping at 124 dB SPL. Tests were made on 25 impaired ears (average losses of 40-90 dB in the range 0.5-2.0 kHz), including 6 with markedly falling audiograms generally diagnosed as sensorineural, and 19 with rising, flat, or gently falling audiograms generally diagnosed as conductive or mixed. For some of the 19 listeners having conductive or mixed losses, broadband "static" noise was added to the speech (independent of the characteristic) to reduce the maximum score to the 50-80% range. The results of the study showed that the characteristic with $n = 0$ or 1 gave the best performance for nearly all the ears tested (and that near the optimum, the differential effects of the various characteristics were not pronounced). On the average, $n = 1$ was slightly better for intelligibility and $n = 0$ for overall "quality" of the amplified speech. For most of the listeners with sensorineural impairments and markedly falling audiograms, $n = 2$ gave scores that were nearly as good as those for $n = 1$, and $n = -1$ produced an overall quality that was nearly as good as that produced by $n = 0$. A supplementary study on bandpass filtering using 12 ears including five with markedly falling audiograms, showed that a frequency range of 500-4000 Hz did not reduce scores for most listeners, and in some cases increased operating range, relative to the 100-7000 Hz range. The authors of the Harvard study concluded that a flat or slightly rising characteristic (that is, $0 \le n \le 1$, or 0-6 dB per octave) over the range 300-4000 Hz would be optimum or near optimum for most listeners. By contrast, only 40% of the listeners achieved best scores with characteristics that mirrored the audiogram. The characteristic suggested by Watson and Knudson, however, would have led to the best scores in roughly 90% of the cases. Note, also, that the Harvard study is consistent with the study of Watson and Knudson in that many listeners with sensorineural impairments preferred characteristics with more low-frequency emphasis (relative to high-frequency) than the characteristics that produced the best intelligibility scores.

Because the Harvard study has had significant impact on the choice of frequency-gain characteristic for use in hearing aids, it is important to mention some objections that have been raised against it. For example, Fletcher (1952) pointed out that the frequency response curves presented ought not be interpreted as functional gain characteristics. For the system with a nominally flat response (n=0), the functional gain was roughly flat in the region 250 to 1000 Hz, fell at 6 dB per octave from 1 to 4 kHz, and fell an additional 1 dB from 4 to 8 kHz. By comparison, the characteristic which nominally provided 6 dB/octave of high-frequency emphasis (n=1), had a functional gain that was roughly flat from 500-8000 Hz, but fell 10 dB from 500 to 250 Hz. Similar corrections were suggested for the other characteristics.

Miller (1972) has pointed out that the listeners tested in the Harvard study do not represent a good sample of the current population requiring hearing aids. Only 18 listeners were tested and, of that number, many had losses with large conductive components. Also, the fact that many of the listeners were sophisticated hearing-aid users, combined with the fact that the listeners were not extensively trained with the experimental characteristics (so that long-term accommodation effects of the type noted by Watson and Knudson were precluded), may have distorted the results. In addition, the monosyllabic word tests used in the study have been criticized as insufficiently sensitive to detect differences in performance associated with change in the frequency-gain characteristic, and as poor predictors of performance with materials more like real speech. Finally, these tests employed only male voices and were administered to the listeners with sensorineural impairments only under quiet conditions.

It is also important to note that the listeners in the Harvard study obtained maximum scores at or near the clipping threshold. Many arguments, including the theoretical analysis presented by Radley et al. (1947), suggest that a moderately rising characteristic should be optimum under such conditions, roughly independent of the loss. Impaired listeners were not tested with a system free of clipping or limiting, and only a small number of characteristics were evaluated with compression limiting (which controls peak delivered power without introducing as much nonlinear distortion as clipping).

Hudgins et al. (1948) extended the Harvard study to an experimental wearable aid, constructed according to the design objectives of the Master-Aid study, and two commercially available aids. When calibrated using a 2-cc coupler, the experimental aid and one of the commercial aids (aid A) had a relatively broad frequency response (20 dB down at 200 and 5500 Hz), whereas the other commercial aid (aid B) had a narrow frequency response (20 dB down at 1000 and 3500 Hz). Also, aid A had less harmonic distortion than aid B. Tests were performed on six listeners with average losses of 48-63 dB in the range 0.5 to 2.0 kHz: two with flat audiograms, three with gently falling audiograms, and one with a markedly falling audiogram. Although roughly the same maximum scores were ob-

tained with the master-aid, the experimental aid, and the better of the two commercial aids, the operating ranges of the wearable aids, and in particular the commercial aids, were significantly inferior to the master-aid. The differences in scores were most pronounced at high input levels, where nonlinear distortion is important, and where the compression-limiting used in the master-aid and experimental aid was advantageous. The authors interpreted their results as confirming the Harvard study and suggested that all types of impaired listeners would obtain best performance with a frequency-gain characteristic that was relatively broad and flat.

Sheets and Hedgecock (1949) extended the work of these studies by testing a much larger number of impaired listeners who were all hearing-aid users and who had a wide variety of audiograms (21 flat, 33 gently falling, 20 markedly falling, 15 rising, and 16 trough-shaped). The characteristics studied were three different response patterns of a single commercial aid, used with the listener's own earmolds, corresponding roughly to

$$G(f) = K \cdot f^{2n}$$

with n = 0, 1.2, and 2.4 (flat, 7 dB/octave rise, and 14 dB/octave rise, respectively) over the range 250-3600 Hz. Listeners with flat, rising, or trough-shaped audiograms achieved best scores with n = 0 and worst scores with n = 2.4. Listeners with gently or markedly falling audiograms achieved best scores with n = 1.2 and worst scores with n = 2.4 (with n = 0 falling in between). The subjective preferences of the listeners (Hedgecock, 1949) for the different characteristics were relatively consistent with the objective measurements: roughly equal numbers preferred characteristics with n = 0 or 1.2 and a significantly smaller number preferred n = 2.4. Also, when the listeners were categorized according to their most-comfortable equal-loudness contours rather than their audiograms (Hedgecock, 1949), the majority (93%) had flat or gently sloping contours, and the remainder had trough-shaped contours (none had rising or markedly falling contours). Sheets and Hedgecock interpreted their results as generally supporting the Harvard study: a flat or moderately rising (but not strongly rising) characteristic appeared best for most listeners. They found no major difference between the results obtained with the flat characteristic and that which most closely mirrored the audiogram, and recommended that if the characteristic were fitted on an individual basis it should be based on the most-comfortable equal-loudness contour, particularly for listeners with markedly falling audiograms.

Radley et al. (1947) conducted a master-aid study in England that in many ways paralleled the Harvard study. They sought to determine a single frequency-gain characteristic that would be of use to the majority of impaired listeners. This characteristic was determined both theoretically

and experimentally with relatively consistent results. The experimental study was conducted in two parts. In the first part, 228 listeners having varying degrees and types of impairments were tested with a variety of characteristics using a master-aid. This aid used over-the-ear telephone receivers calibrated with a 3-cc coupler and included a correction for head diffraction effects on acoustic transmission. For each listener, the speech material was presented at the listener's preferred level, as well as above and below this level. Based on the results of these tests, the authors concluded that the optimum characteristic need extend only to 4000 Hz, and that between 750 and 4000 Hz the amplification should be either uniform or slowly increasing (0-6 dB/octave). Below 750 Hz the optimum characteristic depended on the degree of loss. Listeners with average losses greater than 45 dB required a 12 dB/octave decrease in gain below 750 Hz, while those with smaller losses did equally well with uniform amplification. These conclusions bear close resemblance to those of the Harvard study, particularly with respect to the need for uniform (or slightly increasing) amplification above 750-1000 Hz, and the desirability of attenuating low-frequency signals. Resnick (1977) has pointed out, however, that the nominal characteristics reported in the British study were corrected for head-diffraction effects while those in the Harvard study were not. Thus the gain of the nominally flat characteristic in the British study actually increased about 8 dB from 500 to 4000 Hz relative to the comparable characteristic in the Harvard study.

In the second part of the study by Radley et al. (1947), for which detailed individual results are available, clinical tests were conducted on 27 listeners using six different aids: four experimental aids designed in accordance with the above results, a commercially available American-made aid, and each patient's own aid. The authors interpreted the results of these tests as supporting their design since 20 of the 27 listeners achieved best results with one or another of the experimental aids. Viewed more closely, however, the results are somewhat troublesome. Although both the theoretical and experimental studies indicated that the optimum characteristic for female voices should be roughly the same as for male voices, the actual scores for female voices were much lower than for male voices. Fully one-third of the listeners could not achieve a word recognition score of 40% for female voices with the best aid tested, although all could do so for male voices. Also, many listeners with an average loss greater than 50 dB obtained better scores with their own aids than with any of the experimental aids, in contrast to listeners with smaller losses who did consistently better with the experimental aids.

Knight (1967) redetermined the optimum characteristic when the British master-aid was fitted with insert earphones and equalized to give a uniform response rising 6 dB/octave from 200-4000 Hz. He used tests and levels similar to the earlier study and included 81 listeners with a variety of types of loss. For listeners with sensorineural impairments, better scores were obtained with the uniformly rising characteristic, or with one

that was flat above 1000 Hz, than with a characteristic that had still greater high-frequency emphasis (12 dB/octave above 1000 Hz).

Watson and Tolan (1949) reviewed the results of a wide variety of research on selective amplification, including material obtained from workers out in the field. They noted that about three-fourths of hearing-aid users achieved best scores with a flat gain characteristic when tested with Harvard PB-50 lists presented in quiet using high-quality equipment. However, a characteristic with moderate high-frequency emphasis (3-9 dB/octave) is preferable under a variety of adverse conditions, such as nonlinear distortion, clipping, or environmental noise. They also noted that such a characteristic is best for listeners with sharply falling audiograms, and that, in general, the amount of emphasis required by a given impaired listener is more reliably indicated by an equal-loudness contour well above threshold than by the threshold itself. In addition, they noted that the results obtained by permitting the subject to select his own characteristic are inconsistent. For example, whereas Watson and Knudson (1940) found the subjective choice to be unreliable, Hedgecock (1949) found that a majority of listeners obtained best results with the preferred characteristic for both spondees and monosyllabic words. Finally, Watson and Tolan attribute the conviction of hearing-aid technicians that selective amplification is immensely beneficial for most impaired listeners to the large changes in speech intelligibility, operating range, and comfort that are associated with slight changes in the location and size of the frequency-response irregularities that occur in commercially available aids.

Reddell and Calvert (1966) compared the performance of hearing aids selected according to customary audiological procedures with custom-fitted commercial aids. The adjustment of the commercial aids, which was performed by the manufacturer based on data supplied by the experimenters, was directed at mirroring the audiogram. The study included 24 listeners with sensorineural losses characterized by sloping audiograms (average slope of 18 dB/octave between 500 and 3000 Hz). On the whole, the customized aids (the identity of which was unknown to the subjects) were preferred by the subjects and were slightly superior to the other aids with respect to the SRT, the recognition of monosyllabic words (both in quiet and noise), and tolerance for intense speech. However, as the authors point out, the adjustments made to customize the aids were performed very crudely and did not always coincide with the adjustments requested. Also, no attention was paid in the analyses of the results to other electroacoustic properties of the aids, such as nonlinear distortion.

Katakura et al. (1968) compared flat amplification (within ±2 dB from 200-7000 Hz) to selective amplification chosen to mirror the audiogram for a group of 43 listeners with sensorineural impairments. Selective amplification yielded better maximum scores for 20 of the listeners. Also, of the 16 listeners who neither had flat audiograms nor could achieve scores greater than 80% with the flat characteristic, 15 obtained better results

with selective amplification. (The scores achieved with selective amplification by those whose scores exceeded 80% with the flat characteristic were not reported, however.)

Thompson and Lassman (1969) explored the hypothesis that the advantages of selective amplification are restricted to listeners with low internal distortion. They administered word recognition (W-22) and speech sound comparison (SSC) tests in both quiet and noise together with a variety of clinical distortion tests to a group of 30 listeners with losses characterized by markedly falling audiograms (the average audiogram showed a 2 dB loss at 250 Hz and a 63 dB loss at 4000 Hz). Speech testing, at levels judged best by each subject, employed both amplification with a flat characteristic and amplification that mirrored the group's average audiogram from 250-4000 Hz (TDH-39 earphones, coupler calibrated). The distortion tests, which focused on adaptation and recruitment (Bekésy adaptation, loudness tracking, and threshold width; tone-decay test; and SISI test) were used to divide the listeners into subgroups with similar relative distortion ratings. In general, the results of the speech tests showed that selective amplification is slightly superior to flat amplification. Averaged over the distortion subgroups, the W-22 score increased from 83 to 84% in quiet and from 61 to 65% in noise (S/N = 10 dB); and the SSC score increased from 78 to 83%. The differences between the scores for selective and flat amplification ranged from -1.8 to $\pm 10.8\%$ depending on the subgroup and test, but roughly 85% of the comparisons indicated higher mean scores for selective amplification. Thompson and Lassman interpreted their results as consistent with the hypothesis that selective amplification is superior for low-distortion subjects, but inconsistent with the hypothesis that flat amplification is superior for high-distortion subjects. They also interpreted their results as providing support for the idea that standard speech intelligibility tests (monosyllable recognition in quiet) are an inadequate tool for demonstrating perceptual differences among different amplifying systems (for example, Shore, Bilger, and Hirsh, 1960; Zerlin, 1962; Resnick and Becker, 1963; Chial and Hayes, 1974).

Kemker (1972) compared five frequency-gain characteristics on three groups of subjects with sensorineural losses, one group having flat audiograms, one having gradually sloping audiograms, and one having steeply sloping audiograms. Three of the characteristics were based on perceptual measurements: G_1, derived from the most comfortable equal-loudness contour; G_2, derived from the most comfortable loudness level; and G_3, derived from the slope of the audiogram. Two of the characteristics were independent of perceptual measurements: G_4, a characteristic with a nominal rise of 6 dB/octave; and G_5 a characteristic that was nominally flat. These characteristics were evaluated at sensation levels ranging from 6 to 36 dB and tests were conducted at a signal-to-noise ratio of 6 dB. For all three groups of subjects, characteristics G_1, G_2, and G_4 generally produced scores that were roughly equivalent and superior to those produced by G_3

and G_5. As Kemker points out, the fact that G_1 and G_2 produce similar scores is not surprising: the two characteristics are very similar. Also, and as one might expect, whereas the relative deficiencies of G_3 become apparent only at high sensation levels (and were extreme at these levels for the groups with sloping audiograms), the relative deficiencies of G_5 were either roughly uniform over level or decreased at high levels. For the groups with sloping audiograms, G_3 produced scores at 36 dB SL that were roughly 20 percentage points below the average score produced by G_1, G_2, and G_4. For the group with sharply sloping audiograms, G_5 produced scores at 6, 16, and 26 dB SL that were roughly 15 percentage points below the average score produced by G_1, G_2, and G_4. Kemker concludes that G_1 and G_2 deserve further attention; however, his data do not demonstrate a clear superiority for these characteristics over characteristic G_4 (which does not require individual fitting).

Barfod (1972) performed experiments to determine whether properly chosen high-frequency emphasis could improve speech intelligibility for subjects with sharply sloping sensorineural losses. The average audiogram of the 15 subjects showed a loss of less than 12 dB in the region 125-1000 Hz, a loss of approximately 60 dB at 2 kHz, and a loss of roughly 70 dB at 4 and 8 kHz. The test materials consisted of nonsense syllables mixed with speech-spectrum noise (S/N = 5 dB). In all tests the unprocessed level of the speech was fixed at 60 dB SPL. The nominal frequency-gain characteristics tested (measured at the input to Beyer DT-48 headphones) were flat with a gain of unity up to F Hz, then increased with a slope of S dB/octave to a value A, and then remained flat at all higher frequencies. In the main series of experiments, S was fixed at 24 dB/octave, A varied from 0 to 68 dB, and F varied from 250 to 1500 Hz. The results indicate that, independent of A, the highest intelligibility score is obtained with F = 750 Hz. For one group of subjects, the average scores decreased from 37 to 25% as F increased from 750 to 1500 Hz; for a second group, the scores decreased from 30 to 17% as F decreased from 750 to 250 Hz. For F = 750 Hz and S = 24 dB/octave, the average scores for the two groups were highest (approximately 40%) when A was in the region 40-60 dB (for A = 20 dB, the scores dropped by roughly 8 points). The best score for the first group with the nominally flat characteristic (31% achieved at an overall gain of 40 dB—the maximum gain tested) was 19 points below the score with the best high-frequency emphasis characteristic tested (F = 750 Hz, S = 24 dB/octave, A = 60 dB). Roughly 60% of the subjects found the speech amplified 40 dB by the flat characteristic to be uncomfortably loud; no such reports were provided for the sloping characteristics even though they delivered roughly comparable sound pressure levels. Results of further experiments aimed at determining the effect of the slope S and lowpass filtering at 4 kHz were inconclusive. In general, in interpreting Barfod's results, it should be stressed that the frequency-gain characteristics were measured electrically, not functionally. According to our estimates, maximum intelligibility was achieved with a functional high-

frequency emphasis of roughly 28-48 dB (as opposed to the nominal gain of 40-60 dB). Also, in all experiments the presentation level of low-frequency speech elements was fixed and only the high-frequency gain was varied. Had the presentation level of the low-frequency elements been varied, the amount of high-frequency emphasis suggested would probably also have varied.

Johansen and Simonsen (1973) determined the level dependence of the frequency-gain characteristic judged "most-comfortable" for speech listening. They studied 10 listeners with sensorineural impairments who had essentially no loss at low frequencies but severe losses and recruitment at high frequencies. The characteristics tested were flat up to 1000 Hz, then increased with a slope S to a value P above the low-frequency level, and then remained flat. The input speech level varied from 55 to 95 dB SPL, and the values of the slope S and "treble amplification" P explored were 0, 6, 12, and 24 dB/octave and 0, 12, 20, 30, and 40 dB, respectively. The results indicated (as one would expect because of the known tendency for equal-loudness contours to flatten as the level is increased) that the preferred slope and treble amplification decreases as the input level increases. In a further test, it was shown that none of the listeners found a smoothly rising characteristic with a slope of 6 dB/octave (as recommended by Davis et al.) to be the most comfortable.

Pascoe, Miller, and Niemoeller in carefully controlled studies at the Central Institute for the Deaf using a binaural master hearing aid with insert earphones and headworn miniature microphones, have raised serious questions about the interpretation of much previous research on the frequency-gain characteristic. They have argued that the characteristics specified in many previous experiments are merely nominal and do not take proper account of the acoustic transformations that actually occur as sound passes from a source to the eardrum. The actual gain characteristics of amplification systems that ignore the field-to-eardrum transfer function (including the effects of concha and ear-canal resonances and of head and pinna diffraction) may differ from the nominal characteristics by signficant amounts, particularly when insert receivers calibrated on 2-cc couplers are used. To overcome these difficulties they have determined the actual transfer functions carefully and specified the *functional gain* of the characteristics they have examined by measuring unaided and aided absolute *field* thresholds. They have also criticized many previous studies for ignoring the possibility that frequency components above 4000 Hz are important for speech perception by impaired listeners. In addition, they have suggested that evaluations of frequency-gain characteristics based on tests using phonetically balanced word lists spoken by male voice in quiet environments are likely to be very weak indicators of their effectiveness. Rather, they advocate using both male and female speakers, noisy and reverberant environments as well as quiet ones, and tests using word lists loaded with phonemes that are particularly difficult for impaired listeners. Finally, they have recommended the use of more homogeneous groups of

impaired listeners than have generally been used in past research.

In a preliminary study, Pascoe, Niemoeller, and Miller (1974) tested a single presbycusic listener who had a binaural symmetric loss characterized by a gently sloping audiogram and a hearing loss of 56 dB using a reference characteristic for which the functional gain was nearly constant over the region 100-6300 Hz. [In terms of the listener's loss L(f), the functional gain G(f) was approximately $L(f)^{0.6}$.] When this characteristic was measured with the microphone in a free field and the receiver mounted in a 2-cc coupler, and the difference between coupler and real-ear characteristics taken into account by means of the Sachs-Burkhard data (1972), the gain function was flat from 1000 to 1500 Hz, rose 19 dB between 1500 and 3000 Hz, and then remained flat to 6300 Hz. (This result provides a dramatic illustration of the difference between coupler-specified gain and functional gain.) All other characteristics tested, many of which were chosen on the basis of this coupler-determined reference curve, were specified in the same manner. The tests were conducted with PB monosyllables presented in both quiet and noise (live female voice in a sound-absorbent room, noise spectrum matched to speech spectrum, signal-to-noise ratio = 6 dB). The results showed that, with few exceptions, the reference characteristic was superior to other characteristics in both quiet and noise. In particular, speech discrimination scores decreased significantly as amplification was decreased in the region 1500-6300 Hz by lowering the high-frequency cutoff, or by raising the frequency at which the 19-dB rise began, or by reducing the amplification in the 3000-6300 Hz region. When the reference characteristic was replaced by a flat characteristic (coupler determined), the discrimination scores decreased from 80 to 62% in quiet and from 77 to 49% in noise. Furthermore, the reference characteristic produced substantially better scores than the listener's own aid (which provided much less amplification both below 700 Hz and above 2000 Hz). The only modifications of the reference characteristic that did not seriously degrade discrimination performance were those involving reduced amplification at low frequencies. For example, when the low-frequency cutoff was raised from 100 Hz to 500 Hz, the score increased by four points in quiet and decreased by four points in noise. Similarly, when the reference characteristic was replaced by a characteristic that had a uniform 7 dB/octave rise from 100 to 6300 Hz (coupler determined), the score decreased by only one point in quiet and six points in noise. In general, the results of this study were interpreted as supporting the idea that preservation of the normal head diffraction and resonance effects is important for speech intelligibility in both quiet and noise. Further discussion of this idea, including arguments concerning why these effects might be important for speech intelligibility, has been presented by Miller (1974).

Pascoe (1975) extended this work to a group of eight presbycusic listeners, all of whom had impairments similar to that of the single listener studied earlier, were regular hearing-aid users, and complained of excep-

tional difficulties when listening in noisy environments. The mean unaided binaural field audiogram for the group fell off at roughly 6 dB/octave between 250 and 6300 Hz and showed a loss of roughly 50 dB in the region 1000-2000 Hz. The characteristics studied included UFG (uniform functional gain), UHL (uniform hearing level, in which the field audiogram is mirrored), UCG (uniform gain as measured with a 2-cc coupler), RC6 (coupler specified rise of 6 dB/octave), and AS (commercial aid simulation). In general, the differences between the characteristics as measured functionally and as measured with the coupler were consistent across both characteristics and listeners: coupler measurements underestimated gain at 1200 Hz by roughly 10 dB and overestimated gain at 3000 Hz by roughly 20 dB. In the tests conducted with these characteristics, monosyllabic word-recognition tests recorded by a male and a female speaker were presented through a loudspeaker in a sound absorbent room to a listener directly facing the loudspeaker and with the overall gain adjusted to the most comfortable level. Background noise was composed of three independent speech-spectrum noise sources presented through loudspeakers to the left and right and behind the listener. For the main series of tests, a special word list heavily loaded with difficult phonemic distinctions (referred to as the high-frequency list) was employed as a closed set, and the speech-to-noise ratio (when the noise was present) was 6 dB. Averaged over listeners, talkers, and noise conditions, the mean speech discrimination scores were UHL-73%, RC6-66%, UFG-62%, UCG-56%, and AS-54%. The advantage of UHL over UFG and of RC6 over UCG (11 and 10 points, respectively) reflects the importance of mirroring the audiogram. The advantage of UFG over UCG and of UHL over RC6 (6 and 7 points, respectively) reflects the importance of accounting for acoustic diffraction and resonance effects. On the average, the score in noise was six points lower than the score in quiet, and the score for female voice was 16 points lower than the score for male voice. The minimum difference between male and female voice (7 points) was achieved with the UHL characteristic. In a supplementary test, UHL and AS were compared using both the high-frequency list and standard PB (open set) lists. For the high-frequency list, the advantage of UHL over AS varied from 11 to 24 points, depending upon the talker and noise condition (noise absent or S/N = 0 dB). For the PB lists, the advantage varied from 6 to 22 points. The fact that the advantage was only six points for the PB lists in quiet was cited as further evidence for the fact that such testing does not provide a sensitive tool for distinguishing between amplification systems. In general, the results of Pascoe's study not only underscore the importance of specifying the frequency-gain characteristic functionally, but also strongly suggest that (at least for listeners suffering from presbycusis and having gently sloping audiograms) a characteristic that truly mirrors the audiogram (UHL) is superior to one that provides truly uniform gain (UFG).

Skinner (1976), in a further study at the Central Institute for the Deaf,

explored the extent to which Pascoe's results (in particular, the superiority of mirroring the audiogram) would apply to listeners with noise-induced losses and sharply sloping audiograms. Intelligibility tests were performed in a sound-treated room with the listener facing a loudspeaker and one ear blocked with masking noise presented through an insert receiver (to make the tests monaural). The speech material presented through the loudspeaker channel (which had a bandwidth of roughly 10 kHz) was similar to that used by Pascoe, but tests were conducted only in quiet. The six ears tested had losses that ranged from roughly 0 to 20 dB in the region 125-1000 Hz and from 45 to 90 dB in the region 2-4 kHz. In the region 1-2 kHz the slope of the average audiogram was approximately 60 dB/octave. (None of the subjects, however, were hearing-aid users.) Field threshold and discomfort levels were measured for one-third octave bands of noise, and most comfortable and discomfort levels for the word lists. In the first experiment, the frequency-gain characteristics (synthesized in the channel feeding the loudspeaker) included a uniform (ortho-telephonic) characteristic and four characteristics with high-frequency emphasis fitted individually to each ear tested. Each of the four characteristics mirrored the audiogram, except that the maximum gain was never permitted to exceed 11, 22, 33, or 44 dB, respectively. For each of these five characteristics, intelligibility tests were performed at a variety of levels. The results for the female talker (averaged over listeners) showed that, for each characteristic, the score obtained at the most comfortable level was approximately equal to the score averaged over the intensity range corresponding to conversational speech (50-85 dB SPL), and that this score was approximately 3 points less than the maximum score obtained over all levels tested (73% compared to 76%). The characteristics with the greatest high-frequency emphasis (33 and 44 dB) produced the highest scores at low intensities, but could be tolerated over only a limited dynamic range. The characteristic with the least high-frequency emphasis (uniform) produced the lowest scores at all levels. Based on the mean score over the intensity range corresponding to conversational speech, the best characteristic was the one with 22 dB of high-frequency emphasis. Averaged over levels, this characteristic produced scores that were roughly 8 points higher than those produced by the uniform characteristic. In the second experiment, the 22-dB characteristic was compared to a variety of characteristics derived from it by the addition or deletion of low- or high-frequency gain. In most cases, the subjects performed as well with the 22-dB characteristic as with any of its modifications. The best system (over all experiments) at its optimal level produced scores that were, on the average, 12 points higher than the uniform system at its optimum level (84% compared to 72%).

Although none of the characteristics tested by Skinner actually mirrored the audiogram, the inferior scores obtained with the 44-dB system suggest that mirroring the audiogram would not have been useful for the subjects studied (even for the given "high-frequency" speech material). In this

sense, her results differ from those obtained by Pascoe (a difference that can at least partially be explained by the fact that her subjects had different losses than those of Pascoe). On the other hand, Skinner's results are similar to those of Pascoe in the sense that the best characteristics obtained in the two studies are similar. More specifically, the frequency-gain characteristic obtained by averaging Skinner's average best system and Pascoe's average best system (both normalized to have 0 dB gain at 500 Hz) has a high frequency emphasis of roughly 20 dB and differs by no more than 4 dB from either of the two characteristics from which it was derived. Finally, it should also be noted that Skinner's results are at least grossly consistent with those of Barfod (1972): both studies made use of subjects with sharply sloping losses and both studies found that intelligibility was at a maximum when high-frequency emphasis was employed.

Skinner interpreted her results in terms of the one-third octave band levels of the processed test materials. The relative performance of the different characteristics appeared to be determined by three factors: 1) the audibility of the individual band levels, particularly those from 2.0 to 6.3 kHz; 2) the proximity of the band levels to the threshold of discomfort; and 3) the balance between the low-frequency (0.5-1.0 kHz) and high-frequency (2.0-4.0 kHz) band levels. On the basis of these factors, Skinner suggested that the frequency-gain characteristic should be chosen to amplify speech spoken at conversational levels so that it falls roughly in the middle of the impaired listener's dynamic range. The high-frequency emphasis provided by the characteristic should, however, be restricted so that the imbalance between low- and high-frequency presentation levels does not exceed 15 dB.

D. STUDIES OF HEARING AIDS

In principle, additional information on the choice of frequency-gain characteristic is available in the literature on the selection, fitting, and evaluation of hearing aids (for example, Watson, 1944; Watson and Tolan, 1949; Rubin, 1954; Shore, Bilger, and Hirsh, 1960; Harris, Haines, Kelsey, and Clack, 1961; Resnick and Becker, 1963; Jeffers and Smith, 1964; Jerger, Speaks, and Malmquist, 1966; Jerger, 1967; Miller and Niemoeller, 1967; Castle, 1967; Wallenfels, 1967, 1971; Jerger and Thelin, 1968; Nance and Causey, 1968; Ling, 1971; Gengel, Pascoe, and Shore, 1971; Ross, 1972; Victoreen, 1973a,b; Northern and Downs, 1974; Chial and Hayes, 1974; Millin, 1975; Levitt et al. 1978). However, many of these articles do not contain experimental results, and those that do are only of limited use in determining the effects of frequency-gain characteristic because of the narrow range of characteristics tested, the way in which the characteristics were selected, and the uncontrolled electroacoustic properties of the aids (many of which change as the characteristic is varied). Also, of course, the functional gain characteristic of the aids is subject to the acoustic transmission phenomena discussed previously and is there-

fore often specified incompletely. The following paragraphs review a small but important sample of this literature (see Table 3).

TABLE 3. Studies of hearing aids employing adult listeners.

Study	Date	Listeners[1]	Q/N/C[2]	Voice[3]	Test Materials[4]
Shore, Bilger, and Hirsh	1960	15 I	Q, N	M	PB-50[5]
Harris et al.	1961	20 N, 22 I	Q	M, F	Sentences
Jerger et al.	1966	6 N, 6 I	C	M	Sentences
Jerger and Thelin	1968	15 N, 10 I	C	?	SSI[6]

[1]N = Normal Hearing, I = Impaired Hearing.
[2]Q = Quiet, N = Noise, C = Competing Message.
[3]M = Male Voice, F= Female Voice.
[4]Consult individual studies for test details.
[5]Hirsh et al. (1952).
[6]Speaks and Jerger (1965).

Shore, Bilger, and Hirsh (1960) highlighted the difficulties involved in reliable clinical evaluation of hearing aids. They performed repeated measurements of the SRT and of monosyllabic word-recognition in both quiet and noise, on 15 listeners, five of whom had sensorineural impairments. Four different commercially available aids were used and each was tested with both a "good" and a "bad" tone setting for each listener. In general, they found that the variability of the test results was sufficiently large with respect to any possible underlying differences in the effectiveness of the various aid/tone-setting combinations to preclude reliable choices on the basis of the tests. (The rms deviation of the discriminations scores appears to have been 10% or greater.) The results of this study are clearly important, and have been appropriately influential in the area of clinical evaluation and selection of hearing aids. On the other hand, they are of only limited value for determining the effects of different frequency-gain characteristics. Aside from the fact that in some cases the "good" and "bad" settings led to characteristics that were almost identical, it is unclear how the "good" and "bad" characteristics were related to each listener's impairment. Also, the aids undoubtedly differed with respect to electroacoustic properties other than the frequency-gain characteristic.

Harris, Haines, Kelsey, and Clack (1961) studied the ability of both normal and impaired listeners to understand normal and predistorted colloquial sentences with a variety of commercially available aids. They found that the different aids led to significantly different test scores and that the level of harmonic distortion was a more reliable indicator of aid effectiveness than flatness of frequency response, frequency range, intermodulation distortion, or signal-to-noise ratio. However, as they point out, the choice of aids was not well suited to exploring effects concerned purely with the frequency-gain characteristic.

Jerger, Speaks, and Malmquist (1966) tested both normal and impaired listeners using a multiple-choice sentence-intelligibility task with three aids that differed in frequency-gain characteristic and harmonic distortion (the amounts of harmonic distortion were 4, 11, and 16%). In general, the results of this study showed that the given task was capable of distinguishing among the aids reliably, that the ranking of the aids was roughly independent of the existence and type of impairment, that performance increased strongly with learning, and that the ranking of the aids was dominated (as in the study of Harris et al.) by the amount of harmonic distortion.

Jerger and Thelin (1968) evaluated the correlation between speech understanding and the electroacoustic properties of 21 aids for both normal and impaired listeners using a synthetic sentence identification (SSI) test presented against a competing message background (Speaks, Karmen, and Benitez, 1967). The results for listeners with normal hearing indicated that performance correlated positively with frequency range below 1000 Hz, but negatively with frequency range above 1000 Hz; negatively with frequency-response irregularity; and positively with the amount of harmonic distortion. In attempting to interpret these results, the authors point out that the frequency region of greatest importance for sentence identification is lower than for the identification of monosyllabic words, that the aids with the best high-frequency response tended to have the greatest response irregularity, and that the range of harmonic distortion considered was relatively small. The results for listeners with sensorineural impairments showed that as the audiometric slope increased, the correlation between the performance of normal listeners and the performance of impaired listeners decreased. Impaired listeners with flat audiograms rank-ordered the aids much as normals did. For impaired listeners with sharply falling audiograms, however, none of the electroacoustic properties considered correlated even weakly with the identification test scores. The authors concluded that the SSI task can be used to separate aids behaviorally, that the relationship between electroacoustic properties of aids and speech understanding depends critically on the nature of the speech task, that frequency-response irregularity is a better predictor of performance than the other factors examined, and that differences among aids are at least as important for normal listeners as impaired listeners. Also, since the differences among the aids were found to be at least as important for normal listeners as for impaired listeners, the authors suggested (surprisingly) that future research concerned with the relation between speech understanding and the electroacoustic properties of aids employ normal listeners. Finally, it should be noted that the authors did not take account of the effect that the small, closed, response sets used in their tests were likely to have had on the results that they obtained (see discussion of Speaks, 1967, in Sec. B).

Further information relevant to the frequency-gain characteristic is available in the literature concerned with earmolds (for example, Harford

and Barry, 1965; Lybarger, 1967, 1972; McClellan, 1967; Dodds and Harford, 1968, 1970; Revoile, 1968; Green and Ross, 1968; Green, 1969; Jetty and Rintleman, 1970; Northern and Hattler, 1970; Hodgeson and Murdock, 1970; Epstein, Shoenut, and Sills, 1973). Variations in the earmold, as in other components of hearing aids, can be used to match a given aid to a given listener. On the whole, the main focus of this matching has been to provide high-frequency emphasis for listeners with high-frequency sensorineural losses. Open earmolds tend both to restore natural high-frequency emphasis and to permit low-frequency signals to be heard without amplification. Generally, better speech discrimination results have been reported for open earmolds than standard earmolds in both quiet and noisy backgrounds (for example, Dodds and Harford, Jetty and Rintleman, Hodgeson and Murdock). The acoustic properties of vented earmolds depend on construction details and are quite variable. Experimental results comparing vented earmolds with standard earmolds are rather inconsistent: vented earmolds were found to be superior by McClellan and by Jetty and Rintleman, but not by Dodds and Harford, Revoile, Hodgeson and Murdock, Northern and Hattler, and Epstein, Shoenut, and Sills.

E. STUDIES EMPLOYING CHILDREN

The problem of determining suitable frequency-gain characteristics for hearing-impaired children differs from that for adults in a variety of ways. First, the task of providing adequate compensation for the hearing loss gains increased importance because of the crucial role played by hearing in the overall development of the child and, in particular, the acquisition of speech and language skills. Second, the distribution of types of impairments that is encountered is significantly different; for example, losses due to aging and long-term noise exposure do not occur. Furthermore, many of the children who have been studied have been obtained from schools for the deaf and have very little residual hearing, often confined to very low frequencies (for example, Boothroyd, 1970). As Erber has pointed out (Stark, 1974), many such children have such little residual hearing that amplified acoustic signals can only be detected vibrotactually. In addition, treatment of such losses often involve the use of body-worn aids that eliminate localization cues and have their frequency-gain characteristics modified by such acoustic effects as body-baffling (for example, Nichols et al. 1947). Third, the task of determining the effects of various frequency-gain characteristics is more difficult than with adults because of the obviously increased measurement difficulties and the developmental changes that are likely to occur during any long-term study (in language competence, ability to participate in controlled experiments). Fourth, despite the importance of providing adequate aid to children, the amount of research relevant to the choice of frequency-gain characteristic that has been performed on children is relatively small. In the following

paragraphs, we comment briefly on an important sample of this research (see Table 4). General background information relevant to hearing aids for children is available in Erber (1971), Stark (1974), and Northern and Downs (1974).

TABLE 4. Studies employing children.

Study	Date	Listeners[1]	Age[2]	Q/N[3]	Voice[4]	Test Materials[5]
Boothroyd	1967, 1968	10 N, 25 I	7-15	Q, N	M	CVC Words
Leckie and Ling	1968	12 I	8-17	Q	M, F	"Speech Sounds"
Ling	1969	12 I	?	N	M	Varied
Darbyshire and Reeves	1969	24 I	10-15	?	?	MPVT

[1]N = Normal Hearing, I = Impaired Hearing.
[2]Age in years.
[3]Q = Quiet Background, N = Noise Background.
[4]M = Male Voice, F = Female Voice.
[5]Consult individual studies for details.

The only extensive study of the perception of filtered speech by children was conducted by Boothroyd (1967, 1968), who examined the dependence of speech discrimination on highpass and lowpass filtering (60 dB/octave) in quiet and schoolroom background noise (predominantly low frequency, overall signal-to-noise ratio 2.5 dB). The tests employed laboratory equipment and circumaural earphones and the test stimuli consisted of phonetically balanced lists of CVC monosyllabic words spoken by a male. Responses were scored for correct phoneme identification (which appeared to be less influenced by contextual and vocabulary factors than word identification). Results of these tests with normal-hearing adults were similar to those reported by French and Steinberg (1947), including, for example, a crossover frequency of 1750 Hz. Results with normal-hearing children showed lower scores than with adults for all conditions tested (including the wideband case). The differences were greatest, however, for highpass filtering, a phenomenon reflected in the lower crossover frequency obtained (1500 Hz). Results with 25 children who had sensorineural losses in the range 45-95 dB and audiograms that ranged from relatively flat to sharply falling showed crossover frequencies ranging from 400 to 2000 Hz. For these children, the crossover frequency was highly correlated with the slope of the audiogram, the broadband discrimination score in quiet, and the ratio of the discrimination score in noise to the discrimination score in quiet. In other words, children with a low crossover frequency tended to have sharply falling audiograms, to discriminate poorly in the quiet broadband condition, and to discriminate much more poorly in noise than in quiet. On the other hand, the crossover frequency was not well correlated with the average hearing loss. More than two-thirds of the impaired children tested had crossover frequencies below 1500 Hz and one-third had crossover frequencies below 1000 Hz.

Also, five of the 25 children, who had relatively little loss at low frequencies, achieved slightly higher discrimination scores than normal adult listeners for certain of the lowpass filtering conditions. Finally, one impaired child, who had a sharply falling audiogram, but little loss at low frequencies, obtained substantially improved scores for discrimination in noise when 1000-Hz highpass filtering was introduced. In general, however, the results indicate that most severely impaired children place greater than normal reliance upon low-frequency cues in word identification tasks.

On the whole, the theoretical arguments concerning the selection of frequency-gain characteristic for children are not markedly different from those advanced for adults. However, due primarily to the increased prevalence of losses that are essentially complete except at low frequencies among children in schools for the deaf, the question of extended low-frequency response has received special attention. For example, Ling (1964) has argued that children with little hearing above 500 Hz should benefit from extended low-frequency amplification in situations in which the background noise level is low. Although he did not describe the shape of the desired characteristic in detail, he suggested that the frequency range should extend down to 70 Hz. The value of extended low-frequency response has also been stressed by Asp et al. (1971, 1972).

Boothroyd, in conjunction with his study of filtering, developed three criteria for determining the frequency-gain characteristic for a given listener: audibility, tolerance, and balance. According to this theory, the frequency-gain characteristic should be chosen to make audible those parts of the frequency spectrum that contribute to an individual's speech perception without exceeding the limits of tolerable sounds and without causing the various parts of the amplified spectrum to interfere with, or mask, one another. According to these criteria, the majority of the impaired listeners considered by Boothroyd would require a characteristic with moderate high-frequency emphasis. To illustrate these principles, Boothroyd discussed four typical cases with sloping audiograms. The characteristics suggested generally mirror the audiogram and are described roughly by $G(f) = [L(f)]^{b(f)}$, where $b \approx 1$ at 250 Hz and $b \approx \frac{1}{2}$ at 2000 Hz.

Northern and Downs (1974) suggested that for most children with hearing impairments the frequency-gain characteristic should incorporate high-frequency emphasis, with the amount of emphasis dependent on the slope of the audiogram. They urged that low frequency emphasis should be used judiciously: amplification of low-frequency speech components may aid those with fragmentary low-frequency hearing in controlling voice quality and in supplementing lipreading, but it may also mask the perception of consonant sounds for those with residual hearing at the higher frequencies. More specifically, they recommended that only when there is no residual hearing above 1000 Hz should amplification be extended to frequencies as low as 100 Hz.

The number of empirical studies directly concerned with the choice of

frequency-gain characteristic for children is much smaller than for adults. Ling (1969) and Leckie and Ling (1968) explored the benefits that could be obtained by extending the low-frequency response of hearing aids for children whose residual hearing was confined to very low frequencies. In the 1969 study, 12 children who had losses from birth and whose audiograms showed losses greater than 65 dB at 500 Hz were tested with two classes of aids having different response regions: 300-3500 Hz and 80-3500 Hz. In each case, the aids were worn for a week and then evaluated by tests administered by male voice in a relatively noisy room (55 dB-C scale). (The earlier tests by Leckie and Ling had shown that the audibility of both male and female voices was improved with an extended low-frequency response.) The results indicated that the extended low-frequency response improved the perception of certain suprasegmental features, such as syllabic structure and stress, but not others, such as intonation. At the segmental level, the extended low-frequency response improved the detection of certain voiced phonemes and the identification of vowels within words, but not the identification of consonants.

Darbyshire and Reeves (1969) tested three frequency-gain characteristics on 24 severely impaired children using the Manchester Picture Vocabulary Test at the level judged most comfortable by each child and at levels both above and below this level. The nominal frequency-gain characteristics tested were (a) flat from 70-7000 Hz, (b) rising at roughly 5 dB/octave from 70-7000 Hz and (c) falling at roughly 4 dB/octave from 800-7000 Hz and flat from 70 to 800 Hz. The results showed that children with flat audiograms (losses of 70-90 dB in the region 250-8000 Hz) achieved best scores (86%) with the flat characteristic and second-best scores (70%) with the rising characteristic. Children with sharply falling audiograms (20 dB loss at 250 Hz and 15 dB/octave slope above 250 Hz) achieved best scores (86%) with the rising characteristic and second-best scores (73%) with the flat characteristic. Both of these groups did relatively poorly (57%) with the falling characteristic which emphasized low frequencies. However, children who had essentially no residual hearing except at low frequencies (55 dB loss at 250 Hz and at least 95 dB loss above 2000 Hz) achieved essentially equivalent scores (55% and 58%) with the flat and falling characteristic and much worse scores (37%) with the rising characteristic. Roughly speaking, these results are consistent with the idea that one should amplify by greater amounts in regions of greater loss, except when the magnitude of the loss is very large. It is also interesting to note that the characteristics and levels chosen by the children as most comfortable did not generally coincide with those that produced the best performance on these tests.

Erber (1971) has presented a general review of work on special hearing aids conducted in schools for the deaf, including previously unpublished research conducted at the Central Institute for the Deaf. He points out that the comparative nature of these studies is often compromised by uncontrolled variables, such as the amount and type of out-of-class experi-

ence with the aids, the influence of high levels of ambient noise in class, home, and play environments, and the failure to equate listeners on bases other than average hearing level. Even relatively well-controlled studies appear to yield conflicting results on the value of extended low-frequency amplification. Despite the obvious logic of extending the amplification to those regions where there is some residual hearing, not all results are supportive. One of the obvious factors that limits the actual benefits of low-frequency amplification is the relatively strong low-frequency background noise that is present in many natural environments. A second factor, less obvious and more controversial, is the upward spread of masking.

F. CONCLUDING REMARKS

(1) Although the set of frequency-gain characteristics studied certainly does not exhaust the set of all possible characteristics, it seems clear that there are a large group of listeners with sensorineural impairments who have significant residual hearing but whose speech perception cannot be restored to normal or near-normal by the use of linear amplification, even with the best possible frequency-gain characteristic. Since a sensorineural impairment cannot be modeled as a frequency-dependent attenuation, this result is not surprising.

(2) Despite the effort that has gone into evaluating frequency-gain characteristics, it is extremely difficult to draw any firm conclusions concerning the optimum frequency-gain characteristic for a given listener on the basis of past research.

Much of this research is consistent with the general idea that the optimum characteristic should fall somewhere between a characteristic that is flat and a characteristic that mirrors the audiogram; however, the data are inadequate to determine which of the various rules that have been proposed and that satisfy this constraint are the most useful, how different the characteristics specified by the different rules actually are when applied to a reasonable sample of impaired listeners, and how important these differences are with respect to improving speech perception. Furthermore, there are indications that, at least for some listeners with extreme losses, best results are obtained by amplifying most in regions where the loss is least (that is, by tending to follow the audiogram rather than to mirror it). In general, it seems clear that attempts to specify the best characteristic for speech perception must take account not only of the audiogram, but also of the discomfort level and the characteristics of speech. However, the best method for including these factors, and the extent to which consideration of these factors is sufficient, is uncertain.

Among the factors that contribute to the difficulty of drawing firm conclusions on the basis of past research are the following:

(a) With few exceptions, past studies have not taken adequate account of relevant acoustic effects in their specifications of the

frequency-gain characteristics. Thus, the stated characteristics are often only nominal and have meaning only on a relative basis within each study: the actual characteristics are usually unknown or can be only estimated crudely from the information supplied.

(b) In many cases the results clearly depend on the interaction between the frequency-gain characteristic and other properties of the amplifying system used, such as nonlinear distortion. In particular, the amount of high-frequency emphasis that is desirable probably depends strongly upon the level at which the amplifying system begins to distort, because low-frequency speech components generally have greater amplitude than high-frequency components.

(c) The usefulness of the results is limited by inadequacies of the speech tests used. Among these are such factors as the test-retest variability; the differential and usually uncontrolled effects of practice with different characteristics; and the poorly understood relationship between performance on the given tests and performance in everyday situations (where such factors as visual inputs, knowledge of context, interference, and reverberation are both important and variable). In addition, it is obviously difficult to compare the results of studies that have used different speech materials and testing conditions.

(d) With few exceptions, there has been only very limited characterization of the impairments of the listeners used in the experiments. Not only has there been essentially no effort to provide in-depth psychophysical characterizations of important perceptual features, but the clinical and audiological analyses are often extremely sketchy.

(e) Studies of the perception of filtered speech by impaired listeners have generally not been sufficiently comprehensive or analytic to guide the selection of the frequency-gain characteristic. This is in sharp contrast to the work with normals for whom the effectiveness of speech transmission systems can often be adequately predicted by articulation theory on the basis of the results of filtering studies.

(f) Even if problems of the above type did not exist, the determination of the optimum characteristic for everyday use in the real world would be exceedingly complex because of the wide variety of signals to be understood (for example, speech material, talker, and position of talker relative to listener) and the nature of the acoustic environment (that is, the characteristics of the interference and reverberation). A realistic determination of the optimum characteristic would not only require extensive study of different characteristics in different real-world situations, but also the construction of a probabilistic description of the set of situations in which the characteristic was to be used. Even if the amplification system were adaptive in the sense that different characteristics were available for different situations, the problem of selecting the set of

characteristics to be made available and of devising an efficient method to enable the listener to choose the best characteristic for a given situation would be a formidable one.

(3) In our opinion, the problem of determining optimum frequency-gain characteristics for impaired listeners will change in character as research on other forms of signal processing becomes more advanced. This opinion is based on our belief that although more advanced forms of signal processing will continue to employ some forms of amplification, the dependence of amplification on frequency will be substantially altered in such systems. For example, it is likely that the optimum amount of high-frequency emphasis would be increased for persons with steeply sloping audiograms if appropriate nonlinear processing were available to protect the listener from uncomfortable or painful sounds. Alternately, the required high-frequency emphasis would be reduced if frequency lowering were used to recode high-frequency signals to take better advantage of residual low-frequency hearing.

REFERENCES

AMBROSE, W. R., The effects of frequency bandwidth and harmonic distortion on speech intelligibility for normal and hearing impaired subjects. Proc. 1972 Conference on Speech Communication and Processing, IEEE Cat. No. 72 CHO 596-7 AE, 270-272 (1972).

AMBROSE, W. R., and NEAL, W. R., The effects of frequency bandwidth on speech discrimination by hearing impaired subjects. *J. Aud. Res.*, 13, 224-229 (1973).

ASP, C. W., Measurement of aural speech perception and oral speech production of the hearing impaired. In S. Singh (Ed.), *Measurement Procedures in Speech, Hearing, and Language*. Baltimore: Univ. Park Press, 191-218 (1975).

ASP, C. W., FRENCH, E., LAWSON, G., PARLATO, P., and PETERSON, M. M., A preliminary evaluation of low-frequency amplification for preschool deaf children. Paper presented at the Annual Convention of the American Speech and Hearing Association, Chicago (1971).

ASP, C. W., FRENCH, E. H., BERRY, J. S., and KELLER, J. E., Evaluation of low-frequency amplification, time in therapy, and testing conditions for preschool deaf children. *J. Acoust. Soc. Am.*, 53, 349(A) (1972).

BARFOD, J., Investigations on the optimum corrective frequency response for high-tone hearing loss. The Acoustics Laboratory, Technical Univ. of Denmark, Rpt. No. 4 (1972).

BERANEK, L. L., The design of speech communication systems. *Proc. I.R.E.*, 31, 880-890 (1947).

BLACK, J. W., Equally contributing frequency bands in intelligibility testing. *J. Speech Hearing Res.*, 2, 81-83 (1959).

BOOTHROYD, A., The discrimination by partially hearing children of frequency distorted speech. *Internat. Audiol.*, 6, 136-145 (1967).

BOOTHROYD, A., The selection of hearing aids for children. Unpublished Ph.D thesis, Univ. of Manchester, Manchester, England (1968).

BOOTHROYD, A., Distribution of hearing levels in the student population of the Clarke School for the Deaf. Northampton: Clarke School for the Deaf (1970).

BROOKS, D. N., Gain requirements of hearing aid users. *Scand. Audiol.*, 2, 199-205 (1973).

BURKHARD, M. D., and SACHS, R. M., Anthropometric manikin for acoustic research. *J. Acoust. Soc. Am.*, 58, 214-222 (1975).

BURKHARD, M. D., and SACHS, R. M., Sound pressures in insert earphone couplers and real ears. *J. Speech Hearing Res.*, 20, 799-807 (1977).

BYRNE, D., and FIFIELD, D., Evaluation of hearing aid fittings for infants. *Brit. J. Audiol.*, 8, 47-54 (1974).

BYRNE, D., and TONNISSON, W., Selecting the gain of hearing aids for persons with sensorineural hearing impairments. *Scand. Audiol.*, 5, 51-59 (1976).

CAMPBELL, G. A., Telephonic intelligibility. *Philos. Mag.*, 19, 152-159 (1910).

CASTLE, W. E., Effects of selective narrow-band filtering on the perception by normal listeners of Harvard PB-50 word lists. *J. Speech Hearing Assoc. of Virginia*, **5**, 12-21 (1963).

CASTLE, W. E., Effects of selective filtering on the perception of Harvard PB-50 word lists by listeners with sensorineural hearing losses. *J. Speech Hearing Assoc. of Virginia*, **5**, 16-26 (1964).

CASTLE, W. E., A conference on hearing aid evaluation procedures. ASHA Rpts., No. 2, American Speech and Hearing Association (1967).

CHIAL, M. R., and HAYES, C. S., Hearing aid evaluation methods: some underlying assumptions. *J. Speech Hearing Dis.*, **39**, 270-279 (1974).

COLLARD, J., Calculation of the articulation of a telephone circuit from the circuit constants. *Elec. Comm.*, **8**, 141-163 (1930).

DARBYSHIRE, J., and REEVES, V., The relationship between pure-tone audiograms, frequency response curves, and optimum listening levels—a pilot investigation. *Sound*, **3**, 37-39 (1969).

DAVIS, H., STEVENS, S. S., NICHOLS, R. H., HUDGINS, C. V., PETERSON, G. E., MARQUIS, R. J., and ROSS, D. A., *Hearing Aids—An Experimental Study of Design Objectives.* Cambridge: Harvard Univ. Press (1947).

DE VOS, A. W., The fitting of hearing aids for babies. *Internat. Audiol.*, **7**, 136-141 (1968).

DODDS, E., and HARFORD, E., Modified earpieces and CROS for high-frequency hearing losses. *J. Speech Hearing Res.*, **11**, 204-218 (1968).

DODDS, E., and HARFORD, E., Follow-up report on modified earpieces and CROS for high-frequency hearing losses. *J. Speech Hearing Res.*, **13**, 41-43 (1970).

DUNN, H. K., and FARNSWORTH, D. W., Exploration of pressure field around the human head during speech. *J. Acoust. Soc. Am.*, **19**, 184-199 (1939).

EGAN, J. P., and WIENER, F. M., On the intelligibility of speech in noise. *J. Acoust. Soc. Am.*, **18**, 435-441 (1946).

EGAN, J. P., Articulation testing methods. *Laryngoscope*, **58**, 955-991 (1948).

EPSTEIN, A., SHOENUT, B., and SILLS, J. A., The effect of earmold types on speech discrimination. Paper presented at the 3rd Oticongress, Copenhagen (1973).

ERBER, N. P., Evalution of special hearing aids for deaf children. *J. Speech Hearing Dis.*, **36**, 527-537 (1971).

FAIRBANKS, G., Test of phonemic differentiation: the rhyme test. *J. Acoust. Soc. Am.*, **30**, 596-600 (1958).

FLETCHER, H., The perception of speech sounds by deafened persons. *J. Acoust. Soc. Am.*, **24**, 490-497 (1952).

FLETCHER, H., and GALT, R. H., Perception of speech and its relation to telephony. *J. Acoust. Soc. Am.*, **22**, 89-151 (1950).

FRANKLIN, B., The effect on consonant discrimination of combining a low-frequency passband in one ear with a high-frequency passband in the other ear. *J. Aud. Res.*, **9**, 365-378 (1969).

FRANKLIN, B., The effect of combining low and high-frequency passbands on consonant recognition in the hearing impaired. *J. Speech Hearing Res.*, **18**, 719-727 (1975).

FRENCH, N. R., and STEINBERG, J. C., Factors governing the intelligibility of speech sounds. *J. Acoust. Soc. Am.*, **19**, 90-119 (1947).

GENGEL, R. W., PASCOE, D., and SHORE, I., A frequency-response procedure for evaluating and selecting hearing aids for hearing-impaired children. *J. Speech Hearing Dis.*, **36**, 341-352 (1971).

GIOLAS, T., Comparative intelligibility scores of sentence lists and continuous discourse. *J. Aud. Res.*, **6**, 31-38 (1966).

GIOLAS, T., and EPSTEIN, A., Comparative intelligibility of word lists and continuous discourse. *J. Speech Hearing Res.*, **6**, 349-358 (1963).

GREEN, D. S., Non-occluding earmolds with CROS and IROS hearing aids. *Arch. Otolaryng.*, **89**, 96-106 (1969).

GREEN, D. S., and ROSS, M., The effect of a conventional versus a non-occluding (CROS-type) earmold on the frequency response of a hearing aid. *J. Speech Hearing Res.*, **11**, 638-647 (1968).

GUBERINA, P., *Case studies in the use of restricted bands of frequencies in auditory rehabilitation of deaf.* Office of Vocational Rehabilitation (HEW) Contract OVR-YUGO-2-63, Institute of Phonetics, Faculty of Arts, Zagreb (1972).

HARFORD, E., and BARRY, J., A rehabilitative approach to the problem of unilateral hearing impairment: the contralateral routing of signals (CROS). *J. Speech Hearing Dis.*, **30**, 121-138 (1965).

HARRIS, J. D., HAINES, H. L., KELSEY, P. A., and CLACK, T. D., The relation between speech intelligibility and the electroacoustic characteristics of low fidelity circuitry. *J. Aud. Res.*, 1, 357-381 (1961).

HASKINS, H., A phonetically balanced test of speech discrimination for children. Unpublished M.A. thesis, Northwestern Univ. (1949).

HEDGECOCK, L. D., Prediction of the efficiency of hearing aids from the audiograms. Unpublished Ph.D. thesis, Univ. of Wisconsin (1949).

HIRSH, I. J., DAVIS, H., SILVERMAN, S. R., REYNOLDS, E. G., ELDERT, E., and BENSON, R. W., Development of materials for speech audiometry. *J. Speech Hearing Dis.*, 17, 321-337 (1952).

HIRSH, I. J., REYNOLDS, E. G., and JOSEPH, M., Intelligibility of different speech materials. *J. Acoust. Soc. Am.*, 26, 530-538 (1954).

HODGESON, W. R., and MURDOCK, C., Effect of the earmold on speech intelligibility in hearing aid use. *J. Speech Hearing Res.*, 13, 290-297 (1970).

HUDGINS, C. V., HAWKINS, J. E., KARLIN, J. E., and STEVENS, S. S., The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope*, 1, 57-89 (1947).

HUDGINS, C. V., MARQUIS, R. J., NICHOLS, R. H., PETERSON, G. E., and ROSS, D. A., The comparative performance of an experimental hearing aid and two commercial instruments. *J. Acoust. Soc. Am.*, 20, 241-258 (1948).

HUIZING, H. C., KRUISINGA, R. J. H., and TASELAAR, M., Triplet audiometry: an analysis of band discrimination in speech reception. *Acta Otolaryng.*, 51, 256-259 (1960).

HUIZING, H. C., and REYNTJES, J. A., Recruitment and speech discrimination loss. *Laryngoscope*, 62, 521-527 (1952).

HUIZING, H. C., and TASELAAR, M., Experiments on binaural hearing. *Acta Otolaryng.*, 53, 151-154 (1961).

INGLIS, H. A., Transmission features of new telephone sets. *Bell Syst. Tech. J.*, 17, 358 (1938).

JEFFERS, J., and SMITH, C. R., On hearing aid selection—in part a reply to Resnick and Becker. *Asha*, 9, 504-506 (1964).

JERGER, J., Behavioral correlates of hearing aid performance. *Bull. Prosthetics Res.*, **Spring**, 62-75 (1967).

JERGER, J., SPEAKS, C., and MALMQUIST, C., Hearing aid performance and hearing aid selection. *J. Speech Hearing Res.*, 9, 136-149 (1966).

JERGER, J., and THELIN, J., Effects of electroacoustic characteristics of hearing aids on speech understanding. *Bull. Prosthetics Res.*, Fall, 159-197 (1968).

JETTY, A. J., and RINTLEMAN, W. F., Acoustic coupler effects on speech audiometric scores using a CROS hearing aid. *J. Speech Hearing Res.*, 13, 101-114 (1970).

JOHANSEN, P. A., and SIMONSEN, E., Some examinations to determine the best frequency characteristic of hearing aids for patients with hearing loss in the high-tone range. *Scand. Audiol.*, 2, 55-58 (1973).

KATAKURA, M., YOSHIDA, T., and OKAMOTO, M., Some proposals on realizing the ideal hearing aid. Proc. 6th International Congress on Acoustics, Paper A-4-6, Tokyo (1968).

KEMKER, J. F., The effect of five hearing aid evaluation procedures on speech discrimination. Paper presented at the Annual Convention of the American Speech and Hearing Association, San Francisco (1972).

KILLION, M. C., An experimental wideband hearing aid. *J. Acoust. Soc. Am.*, 59A, S62 (1976).

KLATT, D. H., and STEVENS, K. N., On the automatic recognition of continuous speech: implications from a spectrogram reading experiment. *IEEE Trans. Aud. Electroacoust.*, AU-21, 210-217 (1973).

KNIGHT, J. J., Redetermination of optimum characteristics for a hearing aid with insert earphone. *Internat. Audiol.*, 6, 232-326 (1967).

KRYTER, K. D., Methods for the calculation and use of the articulation index. *J. Acoust. Soc. Am.*, 34, 1689-1697 (1962a).

KRYTER, K. D., Validation of the articulation index. *J. Acoust. Soc. Am.*, 34, 1698-1702 (1962b).

LA BENZ, P. J., The discrimination of filtered speech by normal and hearing-impaired persons. Unpublished Ph.D. thesis, Johns Hopkins Univ. (1953).

LA BENZ, P. J., Potentialities of auditory perception for various levels of hearing loss. Volta Bureau Reprint 683, Washington, D.C. (1956).

LAWRENCE, D. W., and BLACKLEDGE, V. O., Protocol for evaluation of the effect of hearing

aid electroacoustic parameters on perception of amplified speech. *J. Am. Audiol. Soc.*, **2**, 197-201 (1977).

LAWRENCE, D. L., and BYERS, V. W., Identification of voiceless fricatives by high frequency hearing impaired listeners. *J. Speech Hearing Res.*, **12**, 426-434 (1969).

LECKIE, D., and LING, D., Audibility with hearing aids having low frequency characteristics. *Volta Rev.*, **70**, 83-86 (1968).

LEVITT, H., COLLINS, M. J., DUBNO, J. R., RESNICK, S. B., and WHITE, R. E. C., Development of a protocol for the prescriptive fitting of a wearable master hearing aid. CSL Research Report No. 11, C.U.N.Y. Graduate Center, N.Y.C. (1978).

LICKLIDER, J. C. R., and MILLER, G. A., The perception of speech. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology.* New York: John Wiley and Sons (1951).

LINDEN, A., Distorted speech and binaural speech resynthesis tests. *Acta Otolaryng.*, **58**, 32-48 (1964).

LING, D., Implications of hearing aid amplification below 300 cps. *Volta Rev.*, **66**, 723-729 (1964).

LING, D., Speech discrimination by profoundly deaf children using linear and coding amplifiers. *IEEE Trans. Aud. Electroacoust.*, **AU-17**, 298-303 (1969).

LING, D., Conventional hearing aids: an overview. *Volta Rev.*, **September**, 343-381 (1971).

LYBARGER, S. F., Ear mold acoustics. *Audecibel*, **16**, 9-19 (1967).

LYBARGER, S. F., Ear molds. In J. Katz (Ed.), *Handbook of Clinical Audiology.* Baltimore: Williams and Wilkins (1972).

MARTIN, M. C., Hearing aid gain requirements in sensorineural hearing loss. *Brit. J. Audiol.*, **7**, 21-24 (1973).

MARTIN, M. C., and GROVER, B. C., Current hearing-aid performance and the requirements of hearing-impaired persons. *J. Aud. Eng. Soc.*, **24**, 177-181 (1976).

MARTIN, M. C., GROVER, B. C., WORRALL, J. J., and WILLIAMS, V., The effectiveness of hearing aids in a school population. *Brit. J. Audiol.*, **10**, 33-40 (1976).

MCCLELLAN, M. E., Aided speech discrimination in noise with vented and unvented earmolds. *J. Aud. Res.*, **7**, 93-99 (1967).

MILLER, G. A., *Language and Communication.* New York: McGraw-Hill Book Company, Inc. (1951).

MILLER, G. A., and NICELY, P. E., An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Am.*, **27**, 338-352 (1955).

MILLER, J. D., Possible importance of head diffraction and ear-canal resonance for speech perception and hearing-aid design. *J. Acoust. Soc. Am.*, **55**, 462(A) (1974).

MILLER, J. D., and NIEMOELLER, A. F., Hearing aid design and evaluation for a patient with a severe discrimination loss for speech. *J. Speech Hearing Res.*, **10**, 367-372 (1967).

MILLER, M. H., *Hearing Aids.* New York: The Bobbs-Merrill Company (1972).

MILLIN, J. P., Practical and philosophical considerations. In M. Pollack (Ed.), *Amplification for the Hearing Impaired.* New York: Grune and Stratton (1975).

NANCE, G., and CAUSEY, G. D., Hearing-aid evaluation: an examination of two procedures. *Bull. Prosthetics Res.*, **Spring**, 119-124 (1968).

NICHOLS, R. H., MARQUIS, R. J., WIKLUND, W. G., FILLER, A. S., HUDGINS, C. V., and PETERSON, G. E., The influence of body-baffle effects on the performance of hearing aids. *J. Acoust. Soc. Am.*, **19**, 943-951 (1947).

NORTHERN, J. L., and DOWNS, M. P., *Hearing in Children.* Baltimore: Williams and Wilkins (1974).

NORTHERN, J. L., and HATTLER, K. W., Earmold influence on aided speech identification tasks. *J. Speech Hearing Res.*, **13**, 162-172 (1970).

PALVA, A., Filtered speech audiometry. I. Basic studies with Finnish speech toward the creation of a method for the diagnosis of central auditory disorders. *Acta Otolaryng.*, *Suppl.* **210**, (1965).

PARKER, C. D., The effects of the reduction of "short-time fatigue" on speech intelligibility for "perceptively" deafened individuals. Unpublished Ph.D. thesis, Univ. of Iowa (1953).

PASCOE, D. P., Frequency responses of hearing aids and their effects on the speech perception of hearing-impaired subjects. *Ann. Otol. Rhinol. Laryng.*, **84**, Suppl. 23 (1975).

PASCOE, D. P., NIEMOELLER, A. F., and MILLER, J. D., Hearing aid design and evaluation for a presbycusic patient. *J. Acoust. Soc. Am.*, **55**, 461-462 (A) (1974).

PIMONOW, L., The application of synthetic speech to aural rehabilitation. *J. Aud. Res.*, **3**, 73-82 (1963).

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

RADLEY, W. G., BRAGG, W. L., DADSON, R. S., HALLPIKE, C. S., MCMILLAN, D., POCOCK, L. C., and LITTLER, T. S., *Hearing Aids and Audiometers, Report of the Committee on Electroacoustics.* London: His Majesty's Stationery Office (1947).

REDDELL, R. C., and CALVERT, D. R., Selecting a hearing aid by interpreting audiological data. *J. Aud. Res.*, **6**, 445-452 (1966).

RESNICK, S. B., A summary and critical review of the Harvard and Medresco studies on hearing aids. C. S. L. Rpt. No. 10, C.U.N.Y. Graduate Center, New York (1977).

RESNICK, D. M., and BECKER, M., Hearing aid evaluation: a new approach. *Asha*, **8**, 695-699 (1963).

REVOILE, S. G., Speech discrimination ability with ear inserts. *Bull. Prosthetics Res.*, **Fall**, 198-205 (1968).

RHODES, R., Discrimination of filtered CNC lists by normals and hypacusics. *J. Aud. Res.*, **6**, 129-133 (1966).

RICHARDS, D. L., *Telecommunication by Speech.* London: Butterworth and Co. (1973).

ROMANOW, F. F., Methods for measuring the performance of hearing aids. *J. Acoust. Soc. Am.*, **13**, 294-304 (1942).

ROSENTHAL, R. D., LANG, J. K., and LEVITT, H., Speech reception with low-frequency speech energy. *J. Acoust. Soc. Am.*, **57**, 949-955 (1975).

ROSS, M., Hearing Aid Evaluation. In J. Katz (Ed.), *Handbook of Clinical Audiology.* Baltimore: Williams and Wilkins (1972).

RUBIN, H. A., A study of the response to differential amplification by hyperacusics with flat perceptive losses. Unpublished S.M. thesis, Brooklyn College (1954).

SHAW, E. A. G., Ear canal pressure generated by circumaural and supraaural earphones. *J. Acoust. Soc. Am.*, **39**, 471-479 (1966).

SHAW, E. A. G., The external ear. In W. D. Keidel and W. D. Neff (Eds.), *Handbook of Sensory Physiology*, Vol. V/1. Berlin: Springer-Verlag (1973).

SHAW, E. A. G., Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J. Acoust. Soc. Am.*, **56**, 1848-1861 (1974).

SHEETS, B. V., and HEDGECOCK, L. D., Hearing aid amplification for optimum speech reproduction. *J. Speech Hearing Dis.*, **14**, 373-379 (1949).

SHORE, I., BILGER, R. C., and HIRSH, I., Hearing aid evaluation: reliability of repeated measurements. *J. Speech Hearing Dis.*, **25**, 152-167 (1960).

SKINNER, M. W., Speech intelligibility in noise-induced hearing loss: effects of high-frequency compensation. Unpublished Ph.D. thesis, Washington Univ., St. Louis (1976).

SMALDINO, J. J., BURKHARD, M. D., and CHIAL, M. R., The development and use of an anthropometric manikin for hearing aid research. *Audiology*, Nov. (1977).

SPEAKS, C., Intelligibility of filtered synthetic sentences. *J. Speech Hearing Res.*, **10**, 289-298 (1967).

SPEAKS, C., and JERGER, J., Method for measurement of speech identification. *J. Speech Hearing Res.*, **8**, 185-194 (1965).

SPEAKS, C., KARMEN, J. L., and BENITEZ, L., Effect of a competing message on synthetic sentence identification. *J. Speech Hearing Res.*, **10**, 390-396 (1967).

STARK, R., *Sensory Capabilities of Hearing-Impaired Children.* Baltimore: Univ. Park Press (1974).

THOMAS, I. B., and OHLEY, W. J., Intelligibility enhancement through spectral weighting. Proc. 1972 Conference on Speech Communication and Processing, IEEE Cat. No. 72 CHO 596-7 AE, 360-363 (1972).

THOMAS, I. B., and PFANNEBECKER, G., Effects of spectral weighting of speech in hearing-impaired subjects. *J. Aud. Eng. Soc.*, **22**, 690-694 (1974).

THOMPSON, G., and LASSMAN, F., Relationship of auditory distortion test results to speech discrimination through flat versus selective amplifying systems. *J. Speech Hearing Res.*, **12**, 594-606 (1969).

ULRICH, J. H., An experimental study of the acquisition of information from three types of recorded television presentations. *Speech Monogr.*, **24**, 39-45 (1957).

VICTOREEN, J. A., *A Guide to Applied Otometric Procedures.* Colorado Springs: Vicon Instrument Company (1973a).

VICTOREEN, J. A., *Basic Principles of Otometry.* Springfield, Ill: Charles C Thomas (1973b).

VILLCHUR, E., Free-field calibration of earphones. *J. Acoust. Soc. Am.*, **46**, 1527-1534 (1969).

WALLENFELS, H. G., *Hearing Aids on Prescription.* Springfield, Ill.: Charles C Thomas (1967).

WALLENFELS, H. G., *Hearing Aids for Nerve Deafness.* Springfield, Ill.: Charles C Thomas (1971).

WATSON, L. A., Certain fundamental principles in prescribing and fitting hearing aids. *Laryngoscope*, **54**, 531-558 (1944).

WATSON, N. A., and KNUDSON, V. D., Selective amplification in hearing aids. *J. Acoust. Soc. Am.*, **11**, 406-419 (1940).

WATSON, L., and TOLAN, T., *Hearing Tests and Hearing Instruments.* Baltimore: Williams and Wilkins (1949).

ZERLIN, S., A new approach to hearing-aid selection. *J. Speech Hearing Res.*, **5**, 370-376 (1962).

ZWISLOCKI, J. J., An ear-like coupler for earphone calibration. Rpt. LSC-S-9, Laboratory Sensory Communication, Syracuse Univ. (1971).

Chapter III

PREVIOUS RESEARCH ON AMPLITUDE COMPRESSION

A. PRELIMINARY REMARKS

There are a number of indications that listeners with sensorineural impairments require some form of nonlinear amplitude processing as well as linear amplification. Many of these listeners exhibit increased detection thresholds without correspondingly increased "saturation" or "discomfort" thresholds. Thus, when amplification is used to overcome loss of absolute sensitivity, additional amplitude processing is required to prevent intense sounds from causing pain or discomfort. Also, for some listeners with sensorineural impairments, speech intelligibility rises to a maximum and then falls as intensity is increased. Thus, to achieve best speech perception, the overall gain of the amplifying system used by such listeners must be controlled to assure that the proper output level is maintained as the input level varies. Furthermore, many listeners with sensorineural impairments exhibit abnormal loudness functions. Linear amplification can, in general, restore only one equal-loudness contour to normal for such listeners. To correct the entire set of equal-loudness contours, some form of frequency-dependent nonlinear amplitude processing is required. Finally, past research on linear amplification suggests that the shape of the frequency-gain characteristic must be varied as a function of input level to ensure best speech reception. Clearly, nonlinear processing is required to achieve this variation.

There are three types of nonlinear amplitude processing, with distinct level-temporal characteristics, that have been suggested in order to overcome the above limitations of linear amplification. One type, *limiting*, is used to protect the ear from high or painful peak sound levels. Limiters should not affect sounds below a critical level. They must act very rapidly in order to deal with intense sound bursts having rapid onsets. A second type, *automatic volume control* (AVC), is used to keep the long-term average presentation level (measured over time intervals corresponding to phrases or sentences) near that corresponding to maximum intelligibility. Since the action of automatic volume control should be slow compared to the sound variations that occur within syllables, it can be regarded as a

slowly varying linear amplifier. A final type of amplitude processing, which we shall call *"syllabic compression,"* is concerned with altering the short-term intensity relations among speech elements to improve intelligibility. Syllabic compressors must act rapidly enough to respond to level variations associated with transitions between speech sounds. The level characteristics of syllabic compressors are typically chosen to match the range of speech amplitudes to the residual dynamic range of impaired listeners.

Considerable confusion presently exists in the technical and commercial literature concerned with the use of nonlinear amplitude processing in aids for the impaired. Often, attempts are made to design a single processor, with one fixed level-temporal characteristic, to perform two or three of the above functions. In general, this is not possible. For example, the slow gain variation required for automatic volume control is inappropriate for limiting or syllabic compression. Also, the level characteristics of automatic volume control, which are keyed to the average intensity of speech, may conflict with those of a limiter, which are based on peak levels, and with those of syllabic compression, which are based on matching intensity ranges. It is, however, possible to incorporate these three separate functions in a single aid, provided they are arranged in cascade as shown in Figure 1.
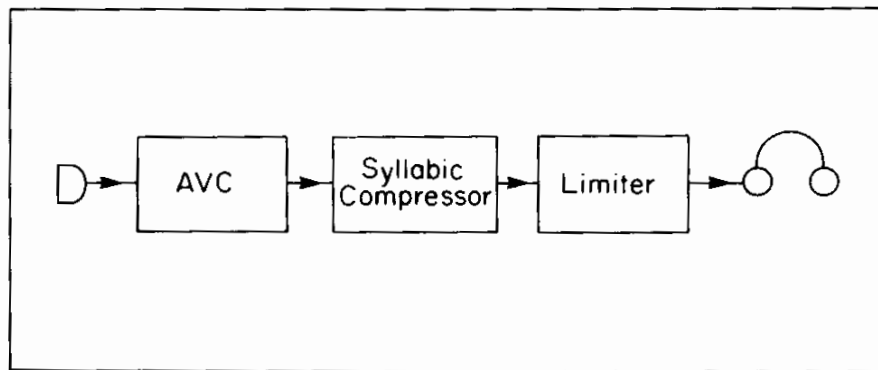
FIGURE 1. Black-box diagram of hearing aid that incorporates three types of nonlinear amplitude processing.

Our review of research on amplitude compression for hearing aids focuses on syllabic compression, the effects of which are least understood and constitute a focal point of current hearing aid research. Because the three types of processing are so often confused, however, we include brief sections concerned with the properties of limiters and automatic volume controls. Also, we preface this review with a discussion of the general characteristics of the amplifiers typically included in compression sys-

tems. Previous reviews concerned with the use of amplitude compression for the hearing impaired are available in Rintlemann (1972) and Villchur (1973).

Our discussion is divided into three sections: (B) Characteristics of Compression Amplifiers, (C) Review of Syllabic Compression, and (D) Concluding Remarks.

B. CHARACTERISTICS OF COMPRESSION AMPLIFIERS

1. *General Characteristics*

The purpose of a compression amplifier is to reduce the dynamic range of a class of input signals. A diagram of the type of system that we will refer to as a "compressor" or "compression amplifier" is presented in Figure 2. In this system the signal is amplified by a low-distortion variable-gain linear amplifier. The gain of the amplifier is controlled by a level detector which senses the input or output signals, or both. Because amplifier gain depends on signal level, the characteristics of the compression amplifier differ from those of a conventional linear amplifier. In particular, although both a compression amplifier and a linear amplifier transform a given sinusoidal input into a sinusoidal output (of the same frequency), the relation of the output amplitude to the input amplitude is different in the two cases. More specifically, if the input amplitude of a sinusoid of frequency f is X, the output amplitude for a linear system is of the form $H(f) \cdot X$ [where $H(f)$ is the amplification of the amplifier], whereas the output amplitude for a compressor is of the form $A(f,X)$. In
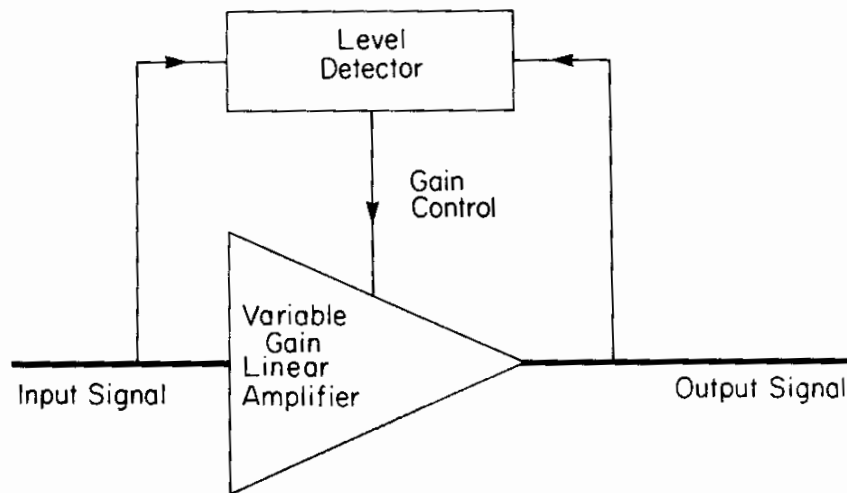

FIGURE 2. Black-box diagram of compression amplifier.

other words, the output of a compressor differs from that of a linear system in that the output amplitude depends nonlinearly on the input amplitude. Ideally, the output of a compressor is free of distortion components and the static gain characteristics of a compressor are described by the function A(f,X). A plot of the output level versus the input level (typically, in dB vs dB) at one frequency describes A(f,X) at that frequency and is called a compression curve (see Figure 3). The *compression ratio* (CR) is the inverse slope of the compression curve (small-range dB change in input divided by the resulting change in output); the *compression threshold* is the input level at which the compression ratio becomes greater than one; the *compression range* is the range of input levels above the compression threshold over which CR > 1 and the output is essentially undistorted; and *expansion* denotes operation in a region on the compression curve where CR < 1. Finally, it should be noted that the phrase "frequency-gain characteristic" has meaning for a compression system only when the input level [X(f)] is specified, and that the frequency-gain characteristic for one input level can be derived from the characteristic for another level by the use of the compression curves. Also, the reference level used to define the frequency-gain characteristic is often chosen (sometimes implicitly) to be the maximum level that one expects to encounter in the use of the system. When a given characteristic is achieved by adding filtering at the output of the compressor, the phrase "post-compression equalization" is often used.
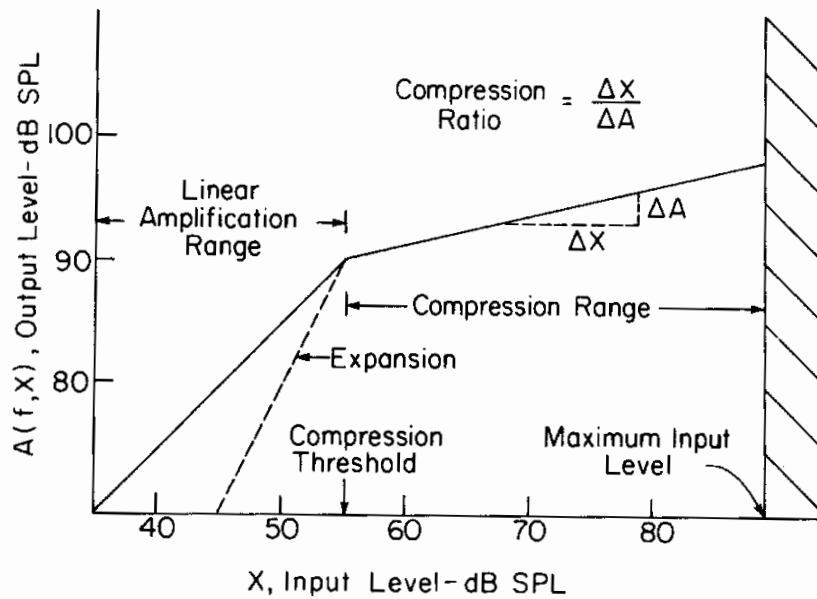

FIGURE 3.   Static compression characteristics.

Static characteristics describe only part of the behavior of a compressor. When the input level of a compressor varies rapidly, as occurs with speech, the dynamic characteristics of a compressor become important. These characteristics are typically described by the *attack time*, TA, and the *release time*, TR, which roughly describe the output envelope after a "step" change in the input envelope (see Figure 4). There are a number of definitions of TA and TR (for example, Carter, 1964; Blesser, 1969; ANSI, 1976). The following are similar to those expressed in ANSI (1976) except the input level change is 20 dB instead of 25 dB. The attack time TA is the time required for the output of a compressor to come within 2 dB of the level specified by the compression curve after the input level increases by at least 20 dB to a level above the compression threshold. The release time TR is the time required for the output of a compressor to come within 2 dB of the level specified by the compression curve after the input level decreases at least 20 dB from a level that was above the compression threshold. These definitions are illustrated in Figure 4. Unfortunately, the time constants TA and TR do not completely describe the dynamic be-
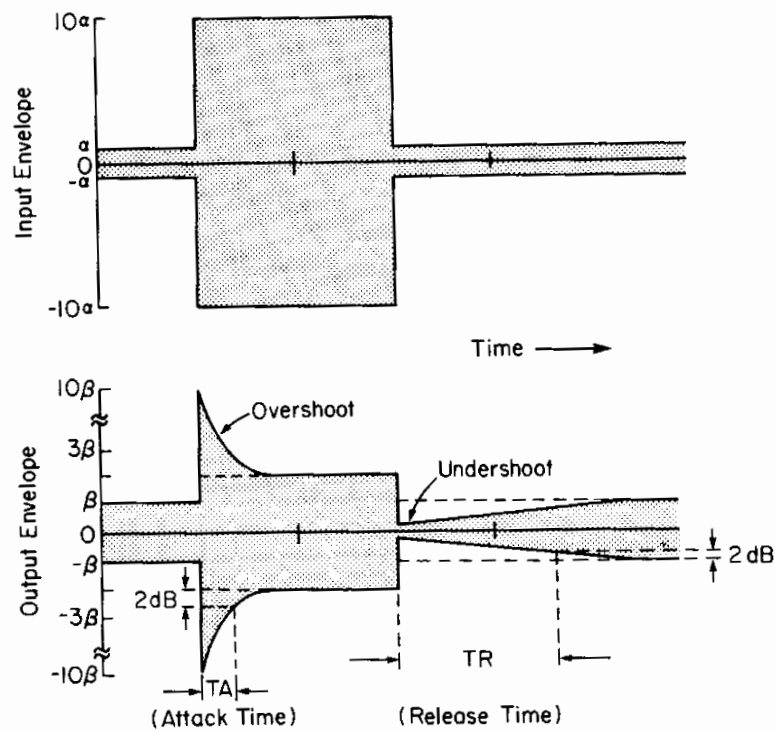

FIGURE 4. Dynamic compression characteristics. The constants α and β are arbitrary.

havior of a compressor. This behavior also depends on the amount and rate of input level change, the input level, the input spectrum, the compression ratio, and the recent history of input level changes. The attack and release time do, however, roughly describe the dynamic behavior of simple compression amplifiers.

There are two fundamental limitations on attack and release times. First, both TA and TR must be much smaller than the time between successive input level changes that are meant to be controlled, so that the static characteristics apply. Second, either TA or TR must be longer than 2-5 periods of the lowest frequency amplified by the system so that compressor action responds to the envelope rather than the instantaneous waveform. If both TA and TR are short, low-frequency components will be distorted (Carter, 1964; Blesser, 1969; Noble and Bird, 1969). Often TA is deliberately chosen to be very short (TA < 1 msec) to prevent excessive overshoot in the output envelope (which might cause discomfort or annoyance) when the input increases rapidly. When this is the case, both of the above limitations can usually be satisfied simultaneously for speech signals by setting TR greater than or equal to 20 msec (for example, Edgardh, 1952; Johansson, 1973; Villchur, 1973). The upper limit on the choice of TR depends, of course, on the intended function of the compressor.

Amplitude processing directed toward the same goals as compression amplification (and often imprecisely referred to as compression amplification in the hearing-aid literature) is sometimes accomplished using a "nonlinear distortion system." Such a system includes instantaneous nonlinear distortion together with pre- and post-distortion filtering. It differs from the compression amplification system shown in Figure 2 in that the response to a sinusoid is generally not a sinusoid (even in the steady state). The nonlinear distortion typically involves signed square rooting, cube rooting, or, as shown in Figure 5, symmetric peak clipping.

The operation of a distortion system is determined by the specific nonlinearity chosen and the pre- and post-distortion filtering used. A compression curve can be plotted for a distortion system (total rms output, or output at input frequency, versus input), but must be interpreted with care because the output signal may be extremely distorted. In multiband distortion systems, out-of-band harmonic and intermodulation distortion components can effectively be eliminated by pre- and post-distortion filtering. However, in-band harmonic and intermodulation components are not eliminated unless the bands are very narrow. The terms attack and release time do not apply to a distortion system.

There is one form of degradation that necessarily accompanies both compression amplification and processing by nonlinear distortion systems: an increase in background noise during periods when the signal is weak or absent. The only sure technique for reducing this degradation, the effects of which are not yet well understood, is to provide a high input signal-to-noise ratio (for example, by appropriate microphone selection
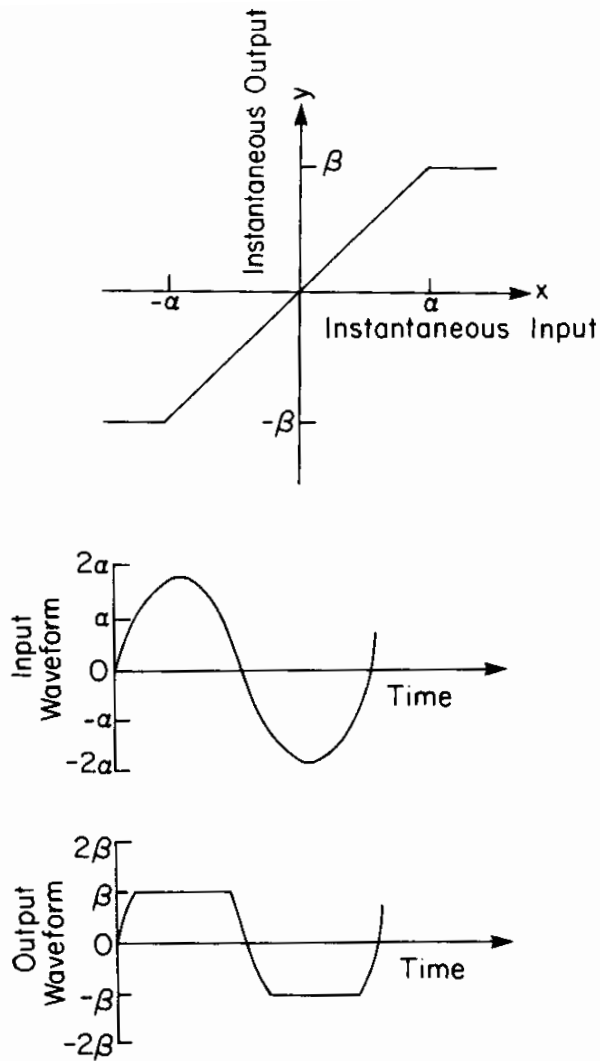
FIGURE 5. Distortion characteristics of an instantaneous peak clipper. The constants α and β are arbitrary.

and placement). In addition, the effect of the noise should always be minimized by such basic design considerations as selecting a compression ratio that is no higher than needed, selecting a compression threshold that is no lower than needed, and selecting a compression curve with a sharp "knee" at the compression threshold. Multichannel compressors provide considerable flexibility in these choices because compression characteris-

tics in each channel can be adjusted to take advantage of spectral variations in signal and noise levels. A number of investigators have also suggested more elaborate techniques, such as modifying the compressor to provide expansion below the compression threshold (Blesser, 1969; Villchur, 1973) or reducing the overall gain when the input is noise alone (Parker, 1953; Hellwarth and Jones, 1967).

Two purely practical problems are encountered when constructing compression hearing aids. First, elements that perform the required functions must be obtained. In the past, the unavailability of such elements has limited the class of compressors that could be constructed economically to those with a high compression ratio (CR > 4). In the last 10 years, however, the introduction of high quality integrated circuits (full-wave rectifiers, rms detectors, log and antilog converters, and multipliers) has overcome this limitation. Second, the compressor must be incorporated in a low-cost portable hearing aid without compromising the desired characteristics. A number of unnecessary distortions in commercial aids (Krebs, 1972; Nabelek, 1973; Burnett and Bassin, 1976; Burnett and Schweitzer, 1977) reflect the failure to solve this problem adequately. These distortions include: "thump," a low-frequency transient that accompanies gain changes and can cause extreme distortion (Hathaway, 1950); harmonic distortion associated with nonlinearity in the gain-controlling element; distortion at low frequencies caused by improper filtering of the gain-control signal; "holes" in the output caused by excessive gain reduction after an increase in input (Blesser and Kent, 1968); excessive overshoot caused by an attack time that is too long; and excessive noise during quiet periods caused by improper adjustment of the compression threshold or compression ratio, or by noise introduced by the gain-controlling device. Many of these unnecessary distortions can be particularly detrimental to speech perception because they tend to occur during the level changes that typically accompany transitions between vowels and consonants. Although theories and electronic devices exist that permit these distortions to be reduced or eliminated (Carter, 1964; Dolby, 1967; Blesser and Kent, 1968; Blesser, 1969; Burwen, 1971; Blackmer, 1972), they have been applied to hearing-aid construction only recently. Most commercial compression aids available today (for example, Viet, 1973) appear to be designed to use a minimum number of components rather than to eliminate distortions.

2. Characteristics of Limiting

The purpose of limiting is to prevent the output sound level from becoming dangerous or uncomfortable (for example, Davis et al. 1947; McCandless, 1973). The importance of controlling level has been underscored by Wallenfels (1967) who stressed that aids which produced uncomfortable sounds were simply not used by impaired listeners. Output sound levels can be controlled by either a peak clipper or a compressor-limiter. A peak clipper distorts the instantaneous waveform, never allow-

ing the instantaneous output to exceed a critical level, as shown in Figure 5. A compressor-limiter rapidly lowers the gain when the input exceeds a critical level and holds the gain down until the input has decreased sufficiently. A compression curve for a compressor-limiter with a maximum output level of about 100 dB SPL is shown in Figure 6. The linear gain of the limiter in Figure 6 is assumed to be unity for signals of amplitude smaller than 100 dB SPL. In practice, the output level at the compression threshold would be adjusted on an individual basis to be somewhat less than the discomfort level. Nominal characteristics of a compressor-limiter are given in the first row of Table 5. Limiters with these characteristics are commonly used in broadcast and recording studios (Shorter et al. 1967; Blesser, 1969; Noble and Bird, 1969). An extensive set of comparisons between clipping and compression techniques for protection against intense sounds is available in the literature, some of which is reviewed below. The characteristics of the compression-limiters included in this review are given in rows 2-5 of Table 5.

Davis et al. (1947), as part of the "Harvard Master Hearing Aid" study, compared clipping to compression-limiting for impaired listeners (sensorineural, conductive, and mixed) by measuring recognition scores for monosyllabic words imbedded in a fixed carrier sentence. With clipping, scores generally decline after the input level is increased beyond the point at which peak levels exceed the clipping threshold, even when highpass filtering is used before clipping. With compression-limiting, no
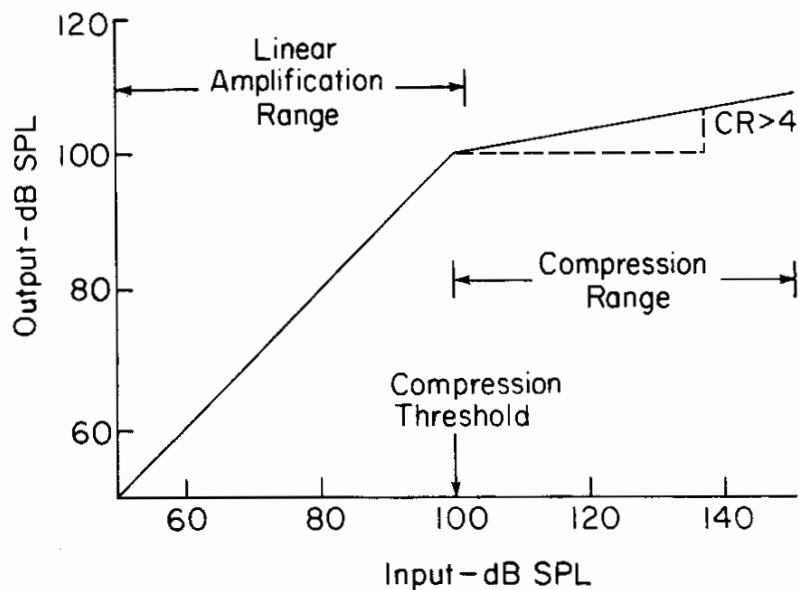
FIGURE 6. Compression curve of a compression-limiter.

TABLE 5. Characteristics of compressor—limiters.

Limiter Type	Attack Time (msec)	Release Time (msec)	Average CR Above Threshold	Output at Compression Threshold (dB SPL 1kHz)	Compression Range (dB)	Distortion Total Harmonic	Distortion Freq (Hz)
Nominal	<1	100 to 1000	>4	100 to 140	20 to 60	<1%	100 to 10k
Davis et al. 1947	1	200	10	117	30	<8%	600 800 1k
Hudgins et al. 1948	15 to 30	50 to 100	3	125	15	<10%	1k
Silverman and Harrison 1951	?	100 to 1000	10	up to 140	25	<2%	?
Blegvad 1974	40	200	3.3	110	40	?	?

such decline is observed; the scores remain essentially constant when the input level is raised above the compression threshold. Hudgins et al. (1948) demonstrated that a wearable hearing aid with compression-limiting could be built and that at high input levels this aid performed better than two commercial aids that simply saturated and thus clipped.

Silverman and Harrison (1951) described a compression-limiter that was used in group hearing aids at schools for the deaf. The limiter freed teachers from worrying about talking too loudly and removed the fear of sudden acoustic shock from the children. A. W. DeVos (1969) commented favorably on a compression-limiter in a group aid that had been used in a school for the deaf since 1954.

Since 1951 a number of researchers have studied the advantages of compression-limiting in commercial hearing aids (for example, Fournier, 1951; Pestalozza, 1953; Portmann and Portmann, 1961; Bizaguet and Veit, 1968). Recently, Blegvad (1974) compared a behind-the-ear aid with compression-limiting to ordinary behind-the-ear aids. Of 42 patients with sensorineural losses who used both types of aids for a period of two months, only 13 preferred the aid with compression-limiting. Similar results were obtained by Ludvigsen and Nielsen (1975) for severely impaired adults and by Brink et al. (1975) for profoundly impaired children.

The recent clinical research discussed above has thus demonstrated no significant advantages for the compression-limiting available in commercial hearing aids. However, it is unclear whether this result is due to intrinsic properties of compression-limiting or to other deficiencies. Some of the inadequacies of commercial compression-aids are illustrated by measurements made at the National Bureau of Standards and reported by Burnett and Bassin (1976) and Burnett and Schweitzer (1977). The characteristics of 81 compression hearing aids were measured using standardized procedures (ANSI, 1976). With volume control on maximum, some of the aids had compression thresholds that were as low as 55 dB SPL in free field. These thresolds are unnecessarily low for most persons with hearing loss and would cause limiting of speech at normal conversational levels, as well as excessive amplification of background noise. Also, some aids produced 112-148 dB SPL output with 80 dB SPL input. Many of these output levels exceed the 100-110 dB SPL limit suggested by McCandless (1973) and by the work of Hood and Poole (1966). For these aids, compression characteristics were probably ineffective in preventing discomfort or pain from intense sounds. In addition, attack times on all aids ranged from 1 to 20 msec and release times from under 10 to more than 500 msec. Very few aids had attack/release times similar to those suggested by Davis et al. (1947) in the "Harvard Master Hearing Aid" study. Finally, some of the aids exhibited unnecessary ringing during output overshoot and undershoot which was similar to that noted by Nabelek (1973). In general, the negative clinical results may be explained by the wide range of characteristics noted in these measurements, by confusion about how these characteristics should be adjusted and the related ab-

sence of individual fitting, and by unnecessary distortions introduced by some of the aids.

3. Characteristics of Automatic Volume Control

Automatic volume control (AVC), which is also referred to as automatic gain control (AGC), adjusts the gain as a function of the long-term average speech input level. An AVC compressor acts very slowly and can thus generally be considered to be a linear amplifier in terms of its effect on the short-term level variations of speech. A compression curve for an AVC compressor is presented in Figure 7. This curve indicates that sounds in the 60-100 dB SPL range would be presented over the 60-65 dB SPL range, and that sounds below 60 dB SPL would be amplified linearly with unity gain. In practice, the gain of an AVC compressor would be adjusted on an individual basis so that signals were presented at levels that led to good speech reception and that were comfortable over the long term. Nominal characteristics of an AVC compressor are given in the first row of Table 6. AVC compressors with these characteristics have long been used in broadcast and recording studios, in portable tape recorders, and as elements of speech processing systems (for example, Kaiser and Bauer, 1962; Hellwarth and Jones, 1967; Torick et al., 1968; Blesser, 1969; Bevan, 1973). Also, hearing aids with these characteristics (specifically TR > 150 msec) have been designated AVC aids by the Swedish Medical Board
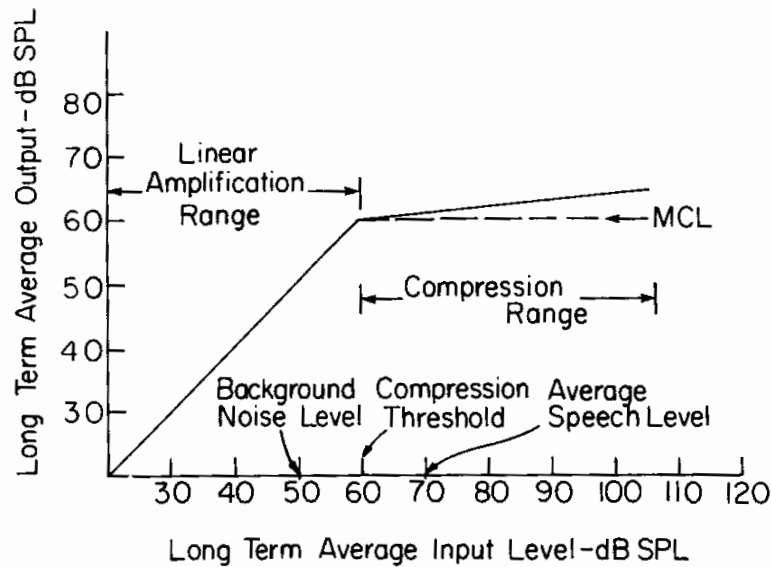
FIGURE 7. Compression curve of an AVC compressor.

TABLE 6. Characteristics of AVC compressors.

AVC Type	Attack Time (sec)	Release Time (sec)	Average CR Above Threshold	Output at Compression Threshold (dB SPL 1kHz)	Compression Range (dB)	Distortion	
						Total Harmonic	Freq (Hz)
Nominal	0.1 to 5	0.15 to 5	>4	near MCL	20 to 60	<1%	100 to 10k
Aspinall	0.01 to 0.1	0.25 to 3	>10	near MCL	30	?	?
Fleming and Rice 1969	0.02	0.2	2, 3, 5	near MCL	?	?	?

(Johansson and Lindblad, 1971). AVC is used in hearing aids to keep the long-term output level near that corresponding to maximum intelligibility while the input level varies (for example, Johansson, 1973). This type of processing might be particularly useful for persons with sensorineural losses who exhibit highly peaked articulation functions (Davis et al., 1947; Huizing and Reyntjes, 1952). Characteristics of some of the AVC compressors that have been studied for possible use with impaired listeners are given in rows 2 and 3 of Table 6.

According to Poliakoff (1950), the first use of AVC was made in a non-wearable aid in 1936. The need for such level control was indicated by a study which showed that out of 500 patients fitted with hearing aids about 50% had a tolerance range for speech of less than 26 dB. Aspinall (1951) described AVC which, together with peak clipping, was incorporated in the "British Master Hearing Aid" to accommodate the reduced tolerance range of listeners with recruitment. Although he suggested that AVC be evaluated with sentences, no results of such tests were given. AVC in conjunction with compression-limiting was incorporated in a group hearing aid designed for use in schools for the deaf (Silverman, 1949), but no evaluation of the effectiveness of AVC in this aid are available.

Flemming and Rice (1969) studied a compressor whose dynamic characteristics were appropriate for AVC, but whose static characteristics were designed to match the dynamic range of speech to the residual hearing of impaired listeners (and hence would probably be more appropriate for syllabic compression). On the basis of a preliminary investigation, involving both normal and impaired listeners, they reported that their compression system was marginally beneficial, but that there was no relation between the optimum compression ratio and the amount of recruitment.

Although there are presently a number of commercial hearing aids that incorporate AVC (for example, Burnett and Schweitzer, 1977), no study has yet demonstrated clear cut advantages of AVC for impaired listeners.

4. *Characteristics of Syllabic Compression*

As noted above, we use the term *syllabic compression* to refer to nonlinear amplitude processing designed to increase speech intelligibility by altering the short-term intensity relations among speech elements. The characteristics of syllabic compressors are often chosen to match the dynamic range of speech to the residual dynamic range of impaired listeners in an attempt to compensate for abnormal loudness function. In general, such processing would tend to restore audibility to low-level speech elements without allowing high-level elements to become abnormally loud. Even if normal loudness relations were restored by syllabic compression, however, normal hearing would not be restored. Not only do sensorineural impairments often result in hearing anomalies other than those associated with loudness (for example, abnormally poor frequency resolution, increased spread of masking, and tinnitus), but amplitude compres-

sion, by its very nature, degrades a listener's effective intensity resolution. All of these factors must be taken into account in the design of syllabic compression systems.

Some insight into the effects of syllabic compression is available in studies which relate speech intelligibility to the consonant-to-vowel ratio (CVR), the ratio of acoustical power in a consonant to that in an adjacent vowel (Fairbanks and Miron, 1957; House et al., 1965; Williams et al., 1966; Salmon, 1970; Hecker, 1974). For syllables in which the consonant is /s/, CVR's ranging from −18 to −9 dB are typically observed, depending on talker and vocal effort. The range of CVR's for other consonants is smaller, but the variation with talker and vocal effort is similar to that for /s/. All studies in which talker or vocal effort were varied have shown that the average CVR is significantly correlated with intelligibility for words presented at equivalent peak levels in a background of additive noise to normal listeners. For example, variation in vocal effort which causes the CVR for /s/ to increase by 6-9 dB is accompanied by an increase of as much as 17 points in the Modified Rhyme Test score (Williams et al., 1966). Similar results for the consonants /ʃ, dʒ, t, tʃ, f, r, w, l/ but not /b/, were obtained by Hecker (1974) who used computer processing to vary the CVR.

A single-channel compression amplifier with appropriate characteristics would achieve increases in the CVR that are roughly similar to those observed in the above studies. Such compression has been shown to improve the intelligibility of speech presented in noise to normals (for example, Kretsinger and Young, 1960), and this suggests that impaired listeners may benefit from such processing, since additive noise simulates many characteristics of recruitment (Stevens, 1966; Richards, 1973).

Theoretically, the benefits of single-channel syllabic compression are likely to be quite limited, however, because the compression curve cannot be varied as a function of frequency and thus compressor action cannot reflect variations in hearing loss with frequency, or changes in the spectral characteristics of the input signal. For example, a single-channel system would not appear to be well matched to listeners with sloping audiograms who exhibit severe recruitment at high frequencies, but normal loudness function at low frequencies. Similarly, a single-channel syllabic compressor is likely to amplify the frication noise of /z/ or /v/ insufficiently because voicing energy controls compressor action for these sounds. To overcome these limitations, a number of investigators have proposed using multichannel syllabic compression systems in which the channels process separate frequency bands and the compression characteristics can thus be made frequency dependent. A simple multichannel system in which each channel compresses a distinct band of frequencies could be used to compensate for reduced dynamic range and recruitment. A more elaborate multichannel system, in which the compressor action for a given band of frequencies is partially controlled by signal components in other bands,

might also be capable of reducing the spread of masking in frequency and time.

Little is presently known about the appropriate choice of static charactersitics for a syllabic compressor. Theoretical arguments based on the properties of speech, the hearing loss, and the environment are still largely in formative stages. By contrast, certain of the dynamic characteristics of syllabic compressors are strongly constrained if the compressor is both to control the level of short speech elements and to introduce minimal distortions in the processed materials. As discussed in Section B-1, the attack time should be less than roughly 1 msec and the release time as short as possible, that is, 20 msec. Not surprisingly, these choices are similar to those suggested more than 30 years ago for use in the speech spectrograph (Dudley and Gruenz, 1946; Steinberg and French, 1946). The same constraints could be applied to multichannel systems, although in such systems the attack time can sometimes be lengthened and the release time correspondingly shortened because filtering reduces rapid onset and offset rates.

C. REVIEW OF SYLLABIC COMPRESSION

In general, studies of syllabic compression have focused on the problem of improving the intelligibility of words or syllables for persons with sensorineural impairments. Very little work has been concerned either with the "quality" of compressed speech or with an analysis of the perceptual confusions that occur with compression. Despite the effort that has been expended, it is still unclear whether syllabic compression is useful in improving speech reception for impaired listeners. In addition to the problems that have been discussed in connection with research on linear amplification (for example, specification of functional gain, inadequacies of speech tests, and limited characterization of impairments), research on syllabic compression has generally been limited by three major problems.

First, syllabic compression systems are usually much more complex than linear amplification systems. As indicated above, the operation of even the simplest compression amplifier depends on the settings of such parameters as attack and release times, compression ratio, compression threshold, compression range, and frequency-gain characteristic. In addition, internal distortion and noise are likely to play a prominent role in compression systems. Because of this complexity (which, of course, increases as the number of channels increases), and because standardization of compression parameters has only recently begun to receive attention, the problem of adequately specifying the systems studied is a difficult one. In most research on compression, the systems considered have not been described adequately. In addition, the complexity of these systems has made it difficult to perform experiments in which the parameters of the system are varied systematically and in which interaction effects are studied.

Second, most studies that have compared syllabic compression to linear amplification have failed to give adequate consideration to the choice of frequency gain characteristic of the linear system. Not only have relevant acoustic effects been ignored in specifying these characteristics, but the characteristics have not been well matched to the listeners used. In most cases, the characteristic has been flat, or nominally flat, independent of the properties of the hearing loss (for example, the shape of the audiogram). As indicated in Chapter II, a flat characteristic is far from optimum for a large fraction of listeners with sensorineural impairments. On the whole, this failure to choose an appropriate frequency-gain characteristic for the linear system tends to exaggerate the benefits associated with amplitude compression. On the other hand, it should also be noted that in most studies very little consideration has been given to the frequency-gain characteristic of the compression system. To what extent these two deficiencies tend to cancel (so that the reported relative performance of compression and linear amplification is roughly correct) is unknown.

Third, there has been inadequate consideration of the selection and variation of speech levels used in the study of syllabic compression. One aspect of this problem concerns the long-term level of the speech material. This parameter is particularly critical for studies of amplitude compression because the signal transformation produced by compression depends on level. Also the effects of varying the input level differ from those of varying the output level (even when the range of output level variation is the same). In addition, when compression is compared to linear amplification, the effect of varying input level is likely to be stronger for the linear system because the gain in the compression system is greater for low-level signals than for high-level signals. Throughout our review of syllabic compression, we tend to focus on intelligibility scores maximized over level rather than on performance at presumably comparable levels. A second aspect of this level problem concerns the variation of levels that occurs within items of a test presented at a given overall level. Within the context of the discussion of different types of compression presented in Section B and the scheme shown in Figure 1, an appropriate test of syllabic compression would incorporate syllable-to-syllable level variations comparable to those encountered in everyday speech, but would employ a fixed, long-term, overall level (theoretically achieved through the use of AVC prior to compression). Unfortunately, most of the tests that have been conducted do not meet this criterion. In particular, the natural intersyllable level variation that occurs in words and phases has often been reduced by artificially preprocessing the test materials (equating levels). In general, it is not possible to estimate the effects of this preprocessing on the intelligibility scores obtained. It seems obvious, however, that this preprocessing reduces the measured effectiveness of syllabic compression relative to that of linear amplification.

The characteristics of the systems considered in this review of syllabic compression (called compression in this section) are given in Table 7.

Edgardh (1952) was among the first to propose using syllabic compression to increase speech intelligibility for impaired listeners. He suggested using a modified form of compression limiting (high compression ratio, low compression threshold, and short release time) to equalize the levels of vowels and consonants for persons with reduced dynamic ranges. Although he does not appear to have evaluated this type of "extreme limitation" with impaired listeners, he reported that processed speech exhibited a certain sibilance, but no "distortion of speech as to affect adversely its comprehension could be observed—either in male or female voices." He further noted, however, that "each breath taken by the talker was amplified to a loud gasp," as would be expected for a system with a high compression ratio and a low compression threshold. Although the high compression ratio (CR > 7) specified by Edgardh has not been recommended by more recent studies, it is noteworthy that Edgardh's suggestions for attack and release times (1-2 msec and 20 msec, respectively) have often been adopted by later investigators.

Parker (1953) experimented with a modified commercial limiter as a means to reduce "short time fatigue" in impaired ears caused by high level sounds. (Results of Parker's study that relate to the effect of highpass filtering are discussed in Chapter II.) He measured the intelligibility of both linearly amplified (nominally flat gain) and compressed PB-50 word lists (Egan, 1948) at presentation levels varying from 6 to 36 dB above the speech reception threshold for linearly amplified spondees. Of the 10 subjects with sensorineural losses studied (average loss roughly 50, 55, and 60 dB at 0.5, 1, and 2 kHz), only four showed an increase in intelligibility with compression (when scores are maximized over level). Averaged over these four subjects, the maximum score obtained with compressed speech exceeded the maximum score obtained with linearly amplified speech by roughly 22 points. However, eight of the 10 subjects (including the above four) also showed an average improvement of roughly 15 points when a sharp highpass filter at 0.67 kHz was applied to the linearly amplified speech (at the input level producing maximum score for the unfiltered speech). Only two of Parker's subjects were benefitted more by compression than by filtering. In interpreting these results it should be noted that the compression system caused increased distortion and noise. It had a high compression ratio and a low compression threshold, "amplified (internal noise during silent periods) to a loudness level that ultimately became almost as high as the speech signal," and had a release time that was short enough to have distorted speech elements with low-frequency components. Also, the above results refer only to scores maximized over level; at low levels, most of the subjects achieved substantially higher scores with compression than with linear amplification.

Kretsinger and Young (1960) used a compressor described by Daniel (1957) and demonstrated that persons with normal hearing listening in noise could achieve improved word intelligibility scores with compression. They compressed over input ranges of 10 and 20 dB, "added white

TABLE 7. Characteristics of syllabic compressors.

Researcher	Compressor Type	Attack Time (msec)	Release Time (msec)	Average CR Above Threshold	Compression Range (dB)	Distortion	
						Total Harmonic	Freq. (Hz)
Edgardh 1952	Single-Channel	1-2	20	10	30	low	0.2-6k
Parker 1953	Single-Channel	<10	10	2.5-3	25-35	(a)	
Kretsinger and Young 1960	Single-Channel Peak Clipping	0.15 NA	22 NA	>10 >10	>20 >20	4.6% 18.0%	0.3-3k 0.3-3k
Lynn and Carhart 1963	Single-Channel	5-85	30-1200	>5	30	?	-
Caraway and Carhart 1967	3-Channel Instantaneous Cube and Square Rooting	NA	NA	2 3	30	13-15% (b) 18-35% (b)	250-500
Trinder 1972	Single-Channel Instantaneous Rooting	NA	NA	>1	>60	>10%	?
Burchfield 1971	Identical to Caraway and Carhart, 1967 Plus Filtering	NA	NA	2 3	30	2-6% (b)	250-500
Vargo 1977	Single-Channel	<1	50	1, 2, 5	<24	?	-
Johansson 1973	Single-Channel	0.5-5	10-1000	?	?	?	-
Thomas and Sparks 1971	Infinite Peak Clipping	NA	NA	>10	?	(c)	

Characteristics of syllabic compressors (cont.)

Researcher	Compressor Type	Attack Time (msec)	Release Time (msec)	Average CR Above Threshold	Compression Range (dB)	Distortion	
						Total Harmonic	Freq. (Hz)
Ruhrberg and Esser 1973	Single-Channel	?	10-20	2.5	47	low	>300
Robinson and Huntington 1973	Single-Channel Time Delay	10-20	(d)	?	?	?	-
Gregory and Drysdale 1976	HFCC (f)	<10	<10	>10	>60	low	0.4-2.5k
Yanick 1973	Single-Channel	1	30	2.6, 2, 1.4	60	?	-
Yanick 1976a	Single-Channel	3	8	2	60	?	-
Nabelek and Robinette 1975	7 Single-Channel Commercial Aids	6-130	30-580	1.5-3	?	(e)	
Villchur 1973	2-Channel	<1	20	>1	>40	<3.5%	0.25-6.5k
Yanick 1976b	1- and 2-Channel	<1	20	>1	40	<3.5%	0.25-6k
Yanick and Drucker 1976	Single-Channel plus Expansion	<1	20	>1	40	<3.5%	0.25-6k
Barfod 1976	1-, 2-, & 3-Channel	6-24	6-24	>1	51	(g)	

a) Silent period noise was almost as loud as the speech signal which was passed through the compressor and re-recorded three times.
b) Distortion = $100\sqrt{(h_2^2 + h_3^2)}/h_1$, where h_i^2 is the energy of the ith harmonic.
c) Only zero crossings preserved.
d) See Figure 7.
e) Distortion during compression overshoot was extreme in some aids and the duration of this distortion ranged from 0 to 70 msec.
f) High-frequency carrier clipping (see text).
g) Distortion was not measured but specifications of system components indicate that total distortion values were similar to those given for Villchur (1973).

noise 3 dB below speech level" to the compressed signals, and presented W-22 lists (Hirsh et al., 1952) at 70 dB SPL to 30 normal hearing subjects. On the average, scores increased from 57% for linear amplification (nominally flat gain) to 85% for 10 dB of compression and 78% for 20 dB of compression. For speech processed by "10 and 20 dB of instantaneous clipping," intelligibility increased to only 63 and 61%.

Lynn and Carhart (1963) investigated the effect of various attack and release times using a compressor "constructed from hearing-aid components." They utilized three groups of listeners (10 per group) with losses that were attributed to otosclerosis, labyrinthine hydrops, and presbycusis. The loss for each group (averaged over the frequencies 0.5, 1, and 2 kHz and over the members of the group) was roughly 35-40 dB. The average audiogram for the otosclerotic group was flat; for the other groups it fell roughly 10 dB per octave above approximately 0.2-0.5 kHz. The investigators first measured the speech reception threshold (SRT) of 30 hearing impaired subjects using processed isolated spondees and found that the SRT decreased by about 10 dB as TA increased from 5 to 85 msec, independent of TR. They then presented paired PB-50 words in carrier sentences (for example, "Please repeat STRIFE BAIT.") 25 dB above the SRT. They found little change in scores for TA/TR ranging from 6/30 to 20/500 msec and generally little improvement in intelligibility for compressed materials. Average discrimination score was constant to within 4 points except for the two extreme values of TA/TR (70/400 and 85/1200 msec) where it decreased by 4 and 12 points. The maximum increase in average score (obtained by first averaging over members of a group and then maximizing over TA/TR) above linear amplification (nominally flat gain) was roughly 9 points (82% vs 73%) for the presbycusis group, 8 points (80% vs 72%) for the hydrops group, and 2 points (93% vs 91%) for the otosclerosis group.

As noted by the authors, the reduction in SRT with increasing TA would be expected for isolated spondees. When TA is large, the beginning of the spondee was amplified with maximum gain to a level that was roughly 45 dB above SRT (because of the initial overshoot pictured in Figure 4). Under this condition, the measured value of SRT would clearly be lower than for short attack times. Since the PB-50 words were presented 25 dB above the measured SRT, the decrease in scores at long TA/TR times may simply reflect reduced presentation level. (Intelligibility scores for monosyllables often increase 2-4% per dB in the region 10-20 dB above SRT.) Also, extreme compression was used and, although distortion was not specified, the compressor used is likely to have been subject to distortions characteristic of commercial aids (for example, Nabelek, 1973). Finally, the 10 subjects with otosclerosis achieved high word scores with linear amplification (greater than 90% on the average) and, to the extent that their losses were purely conductive, would probably not have had a reduced dynamic range requiring compression.

Caraway (1964) and Caraway and Carhart (1967) used a three-channel

instantaneous compressor with square- or cube-rooting. Although the signals in the three bands (0.2-1 kHz, 1-2 kHz, and 2-5 kHz) were processed independently, the same compression ratio (2 or 3) was used for all bands. Compression was compared to linear amplification (nominally flat gain) for CNC words Tillman et al., 1963) in a carrier sentence with words equated for peak power after processing and presented at sensation levels in the range 0-24 dB SL. Scores for normal listeners were essentially perfect for all systems at 24 dB SL. Scores for impaired listeners (labyrinthine otosclerosis, labyrinthine hydrops, and presbycusis; average audiograms similar to those in Lynn and Carhart, 1963) showed only slight improvement over linear amplification. More specifically, the maximum increase in average score (obtained by first averaging over members of a group and then maximizing over presentation level) was less than or equal to 4 points for all patient groups. Also, the advantage of compression at reduced presentation levels was only slightly greater than at higher levels. It should be noted, however, that many of the impaired subjects were able to achieve very high scores with linear amplification: more than a third of the 36 impaired listeners studied achieved scores in the range 90-98% at 24 dB SL, and the average score for all impaired listeners at this level was roughly 80%. For these listeners and for the speech tests used, it would be difficult to demonstrate an improvement in intelligibility with any kind of speech processing. The performance of some of the listeners may also have been limited by certain characteristics of the compression system. For example, the compression ratio was constant over the frequency range of the system rather than matched to the reduction in dynamic range. Also, high levels of harmonic distortion occurred at all input levels and was extreme at low frequencies. Intermodulation distortion, which may have been even larger, was not measured. Similar results for impaired listeners have been reported by Trinder (1972) who used wide-band instantaneous square-, cube-, and Nth-rooting. He found, however, that this processing severely reduced speech intelligibility for normals listening in quiet.

Burchfield (1971) modified the compressor used by Caraway and Carhart to reduce harmonic distortion by filtering. The filtering employed, however, could not eliminate in-band intermodulation distortion. This modified compressor was compared to linear amplification (nominally flat gain) for 36 listeners with sensorineural losses accompanied by recruitment (flat or gently sloping audiograms; loss roughly 50 dB averaged over subjects and frequencies 0.5, 1, and 2 kHz). Intelligibility was measured using CNC words (N.U. Auditory Text #6; Tillman and Carhart, 1966) in a carrier sentence presented at a peak level that was 24 dB above each subject's SRT measured with linear amplification. The average increase in intelligibility obtained with compression over that obtained with linear amplification was roughly 12 points (63% to 74%) for both CR = 2 and CR = 3.

Vargo and Carhart (1972) and Vargo (1977) reported on a series of experiments designed to resolve the difference between the findings of Cara-

way and Carhart (1967) and the more positive findings of Burchfield (1971). However, they used a single-channel compressor whereas the previous investigators had used three-band instantaneous distortion systems. They compared a linear system (nominally flat gain) to the compression system adjusted to have CR values of 2 and 5 using CNC words (Peterson and Lehiste, 1962) equated for peak level after processing. All materials were presented at 10, 20, and 30 dB SL to 12 normal hearing subjects and to 9 subjects with Meniere's disease and relatively flat losses (loss approximately 50 dB averaged over 0.5, 1 and 2 kHz, and over subjects). The differences in mean scores obtained with the 3 systems (linear, CR = 2, CR = 5) did not exceed 3 points for either group of listeners at any presentation level. This result was interpreted as supporting the results of Caraway and Carhart.

Johansson and Lindblad (1971) and Johansson (1973) suggested the values TA = 2 msec, TR = 20 msec, and CR < 6 under quiet listening conditions for subjects with narrow dynamic ranges. Also, based on experience in Swedish schools for the hard of hearing, they suggested using CR < 5 in the presence of classroom noise. In addition, Johansson (1973) reported the results of experiments that investigated the effect of varying the attack and release times in a single-channel compressor on the intelligibility of CVC nonsense syllables. In this test, TR = 200·TA, and TR was varied from 10 to 1000 msec. Ten normal subjects were tested with flat noise added after processing (S/N = 5 dB) and 12 impaired subjects were tested in silence. For both groups, vowel intelligibility was independent of TR, but final-consonant intelligibility fell when TR was greater than roughly 150 msec. For the normals, a release time of 1 sec increased confusions for the weaker voiceless consonants and /m/.

Thomas and Sparks (1971) compared the intelligibility of PB-50 words that were highpass filtered (12 dB per octave at 1.1 kHz) and infinitely clipped with that of speech that was linearly amplified (nominally flat gain) and had the same average rms level (10, 20, 30, and 40 dB SL). They tested 16 impaired listeners with a variety of etiologies and hearing losses (the loss of individual subjects, averaged over 0.5, 1, and 2 kHz, ranged from 15 to 75 dB). Maximum scores with both types of processing were obtained at the highest levels tested and showed little advantage for the filtered-clipped speech (less than 5 points averaged over subjects). A substantial advantage, however, was obtained at the lower levels.

Ruhrberg and Esser (1973) evaluated the intelligibility of both compressed and linearly amplified (nominally flat gain) monosyllables presented at equivalent peak levels to listeners with cochlear and retrocochlear impairments. For the listeners with moderate cochlear impairments, scores with compression averaged 10 points higher; for severe cochlear impairments, the improvement was 27 points. Listeners with retrocochlear losses achieved smaller improvements: 5 and 17 points, respectively.

Robinson and Huntington (1973) evaluated a modification of the compressor shown in Figure 2 in which a time delay is introduced between

the input and the variable-gain linear amplifier. This modified system is similar to a compression-limiter described by Shorter et al. (1967) and was utilized in an attempt to obtain short (several milliseconds or less) time constants. Although this delay tends to compensate for the time required for the level detector to sense changes in the input signal, and to reduce the overshoot and undershoot in the output envelope (see Figure 8), it does not (as suggested by the authors) circumvent the limitations on attack and release time discussed in Section B-1 of this Chapter. Nevertheless, the basic idea of introducing a delay to permit the compression system effectively to look ahead in time is important and deserves serious consideration. Robinson and Huntington measured intelligibility using monosyllabic word lists in quiet and noise with both normal and hearing impaired subjects. Their results indicated that under some conditions the intelligibility of compressed speech is superior to that of linearly amplified speech. Unfortunately, no detailed results of this study have been made available.

Gregory and Drysdale (1976) employed "high frequency carrier clipping" (HFCC) to achieve amplitude compression with short attack and release times. In this scheme, a high frequency carrier is amplitude modulated by speech, clipped, and filtered to remove distortion components. The resulting signal is then heterodyned back to the original audio frequencies. HFCC was compared to a system with instantaneous peak clipping and to a linear system (nominally flat gain). The bandwidths of all systems were restricted to 400-2500 Hz and 16 dB of clipping was used for both the HFCC system and the instantaneous clipper. This 16 dB of clipping was obtained by first measuring the output peak speech level without clipping, increasing the input level by 16 dB, and then adjusting the clipping to return the output to the original level. Words and sentences were presented to children and adults with sensorineural losses in quiet and also to normals in broadband noise (S/N = 20-30 dB). It was found that seven of the 16 impaired listeners tested with words and four of the seven impaired listeners tested with sentences had scores with HFCC that were more than 10 points higher than scores obtained with linear amplification (roughly 50%). Each of the 12 normals tested with words and seven of the eight normals tested with sentences demonstrated increased scores with HFCC. Also, peak clipping was less effective than HFCC.

Yanick (1973), in a study involving 12 listeners with mild to moderate sensorineural losses, compared a custom-fit, wearable, single-channel compression aid having a low compression threshold to each subject's own aid in free field. The gain and compression ratio of the experimental aid was adjusted for each subject on the basis of the SRT and the speech discomfort level. As would be expected, compression significantly increased the range between each subject's SRT and discomfort level for speech. It also increased intelligibility (measured with PB-50 word lists) at input levels of 45 and 70 dB SPL. At the higher level the increase was
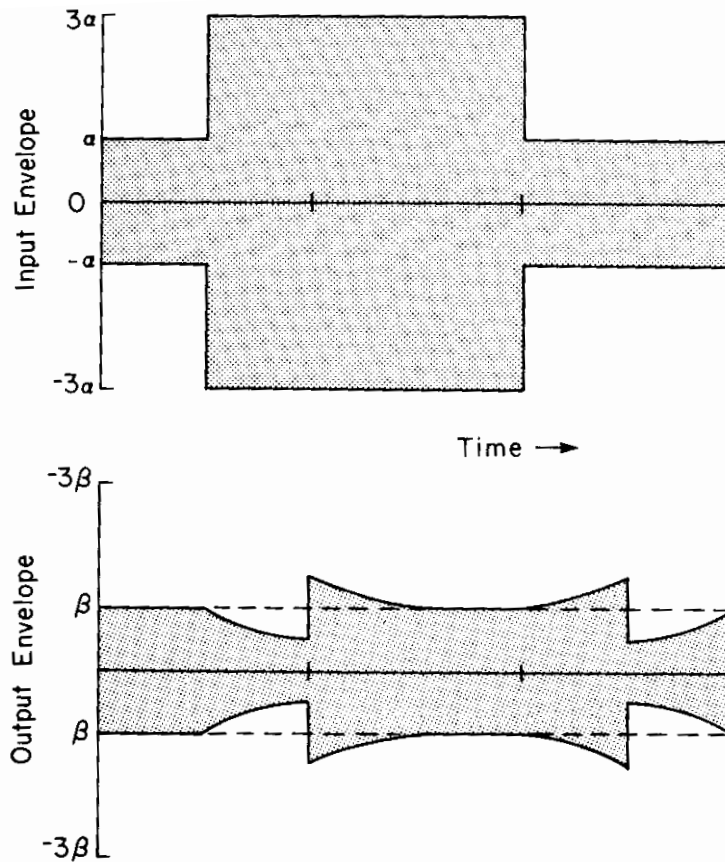
FIGURE 8. Dynamic characteristics of a compressor of the type suggested by Robinson and Huntington (1973) assuming a large compression ratio.

dramatic: 91% for compression compared to 39% for the subject's own aid. The author attributed the advantage of compression to overload of the subject's own aid at the high input level. This work demonstrates that compression can be built into a wearable aid without sacrificing characteristics important for speech perception.

Yanick (1976a) also compared a single-channel, wearable, compression aid having very short time constants to a second wearable aid with compression-limiting. Six subjects with mild to moderate sensorineural losses having flat or gently sloping audiograms received intelligibility tests (PB-50 words) both in quiet and in the presence of speech babble. When the input to the aid with compression-limiting was adjusted such that the compression threshold was not exceeded and the aid was thus linear, the output levels of the two aids were 85 and 90 dB SPL, respec-

tively. Under this condition, the average score for compression-limiting minus the average score for compression was −2, 4, 14, and 0 points at S/N = 0, 10, 20 dB and in quiet. Thus there was little or no advantage demonstrated for compression.

Nabelek and Robinette (1975) evaluated the effect of distortions that occur during compression overshoot in commercial hearing aids on Modified Rhyme Test scores (Kreul et al., 1968). They tested 7 aids using 10 normals listening in noise and 10 listeners with sensorineural impairments (loss approximately 50 dB averaged over 0.5, 1, and 2 kHz and over subjects). In general, poorer scores were obtained with the aids that exhibited greater compression related distortion. Also, the two aids with the shortest attack/release times (6/30 and 17/50 msec) always ranked among the top three in terms of MRT scores. As pointed out by the authors, these results are preliminary: only a limited sample of aids was tested and many characteristics of the aids were not equated or varied independently.

One of the most encouraging studies of syllabic compression was made by Villchur (1973) using high quality commercial (DBX) compressors and specially designed and properly calibrated circumaural earphones. He designed his compression system to "restore normal loudness to each acoustical speech element of importance," noting that this might require many frequency bands, each with its own compressor. At least two bands are necessary because "without multiband compression, only the amplitude ratio between successive speech elements can be changed, and not that between elements that occur simultaneously." His system utilized two compression bands with an adjustable crossover frequency and a ⅓-octave filter bank for post-compression frequency equalization. This design permitted independent adjustment of the compression ratio in the two bands as well as the frequency-gain characteristic of the system as a whole.

Villchur first tested his system with two normal-hearing subjects and a shaped noise background (to simulate hearing loss) that raised thresholds by about 70 dB at 0.5 kHz, 75 dB at 1 kHz, and 90 dB at 2 kHz. Intelligibility was measured using CVC nonsense syllables imbedded in sentences. Materials were recorded by a female speaker in a slightly reverberant environment. Compression ratios were determined from the reduction in dynamic range caused by the noise and equalization was chosen to place the compressed speech at a level within the reduced dynamic range equivalent to the corresponding level in the normal case. Presentation levels were chosen by each subject "on the basis of maximum clarity consistent with long-range comfort." The results show that scores for final consonants increased from roughly 38% with functionally flat linear amplification to 84% with compression and equalization. The average score was 58% with compression alone and 51% with equalization alone. Thus, although compression alone provided some increase in intelligibility, compression combined with equalization provided a much larger increase.

Villchur also measured intelligibility for the same speech material using

six subjects with moderate to severe sensorineural losses (average loss roughly 41, 47, and 67 dB at 0.5, 1, and 2 kHz). Initially he calculated the equalization and compression needed to restore normal loudness relationships for speech on the basis of pure tone threshold, discomfort level, and equal-loudness measurements, as well as the characteristics of speech. On the basis of this calculation, he chose the crossover frequency defining the two bands (this frequency ranged from 1.3-2.5 kHz for different subjects), the compression ratios in the two bands, and the post-compression equalization. Then he allowed subjects to vary frequency equalization and compression ratios from their calculated values while listening to continuous speech to achieve maximum intelligibility consistent with long-range comfort. The adjusted values of the compression ratios were usually near the calculated values: the average values of CR in the low and high bands were 2.2 and 3.5 before adjustment, and 2.1 and 2.8 after adjustment, respectively. The compression systems resulting from these procedures were compared to a linear system that had a functionally flat frequency-gain characteristic except for subject-selected low-frequency rolloff. Averaged over subjects, the linear system had a functional gain that was roughly flat above 500 Hz, but fell at roughly 15 dB/octave below 500 Hz. Speech testing occurred both in quiet and with a competing voice 10 dB below the test words. The speech materials were presented at the most comfortable level (MCL), and with the speech level at the input to the compressor reduced 10 and 20 dB. The results at MCL in quiet show an improvement in terminal consonant scores ranging from 5 points (40 to 45%) to 40 points (25 to 65%) with an average of 21 points, and an improvement in initial consonant recognition ranging from 3 points (60 to 63%) to 22 points (44 to 66%) with an average of 10 points. The high intelligibility of vowels obtained with the linear system at MCL was either maintained or slightly increased by compression. Furthermore, "almost all of the reduced-input scores showed increased benefit from processing, and the improvement from processing was usually maintained or increased in the interference tests." Two subjects were also tested with Harvard sentences (IEEE, 1969). With compression, key word scores increased from 28 to 48% for one subject and from 73 to 89% for the other.

Villchur's results suggest that syllabic compression can be of considerable value. Those elements of his study that appear to contribute to these results include the use of multichannel compression combined with frequency equalization; the determination of compression and equalization parameters on the basis of each subject's hearing characteristics and individual adjustments; the use of equipment with low distortion and noise characteristics; and the use of subjects who potentially could derive large benefits from such processing. It should be noted, however, that the advantage of compression demonstrated in his study may have been artificially inflated by comparison to an inferior linear amplification system (that is, a system with inadequate high-frequency emphasis).

Yanick (1976b) performed a series of experiments that were similar in

design to those of Villchur (1973) but differed in that a background of cafeteria noise was always present, lower S/N ratios were used, and both one-channel and two-channel systems were tested. Yanick's two-channel system was similar to Villchur's. It had a crossover frequency of 1.5 kHz and was individually fitted by allowing subjects to adjust high- and low-frequency channel compression ratios, the gain of the high- relative to the low-frequency channel, and "treble" emphasis and "bass" rolloff (high-pass filtering with a slope of 0, 6, 12, or 18 dB/octave below 1.5 kHz). The resulting system was compared to a linear system that was electrically flat above 1.5 kHz, modified by subject-adjusted bass rolloff similar to that used with the compression system. All adjustments were made while listening to a sentence on a tape loop, and all materials were presented through an insert hearing-aid receiver at the level thought (by the subject) to provide maximum intelligibility. Systems were compared using Harvard sentences that had been recorded by a male talker in a reverberant environment and presented at S/N ratios of 0 and 6 dB. All subjects had moderate to severe sensorineural impairments (losses of 40-70 dB in the "speech frequencies"). One group of 12 subjects had flat audiograms (from 0.5 to 4 kHz); a second group of 12 had sloping audiograms (12-25 dB/octave above 0.75 kHz).

The main experiment compared the linear system to the two-channel compression system. The two-channel compression system had average low- and high-frequency CR values of 2.4 and 2.8 for the flat group and 1.5 and 2.5 for the sloping group, and an average bass rolloff of 6 dB/octave for the flat group and 12 dB/octave for the sloping group. The key word scores showed large gains for the two-channel compression system for both subject groups and both S/N ratios: 37 points (54-91%) for the flat group and 28 points (56-84%) for the sloping group at S/N = 6 dB; 34 points (38-72%) for the flat group and 38 points (29-67%) for the sloping group at S/N = 0 dB. Further experiments, performed using 6 subjects from each of the two groups, demonstrated that performance with each subject's own hearing aid was roughly comparable to that with the linear amplification system. They also showed that a single-channel compression system with an average CR value of 1.8 performed significantly worse than the two-channel system. Scores, averaged over all subjects, for linear amplification, single-channel compression, and two-channel compression were 61, 65, and 91% at S/N = 6 dB and 43, 30, and 78% at S/N = 0 dB.

There are two major aspects of Yanick's study that should be noted in interpreting his results. The first concerns the frequency-gain characteristics of the systems studied. Although the subjects were allowed to adjust many characteristics of the equalization provided in the two-channel compression system, they were only allowed to adjust the bass rolloff of the linear system. These procedures, combined with the use of an insert receiver (which tends to reduce the gain at high frequencies), may have artificially inflated the scores for the two-channel compression system relative to those for the linear system. The second element of Yanick's study

that requires comment concerns the use of low S/N ratios for a noise (cafeteria noise) whose low-frequency components are likely to predominate. Under such conditions, the gain of the single-channel system and of the low-frequency channel of the two-channel system may be controlled primarily by the background noise rather than by the speech signal. The single-channel system would then operate as a linear system with roughly constant gain, and the two-channel system would operate as a single-channel compression system with compression confined to the high-frequency channel. This could explain the poor performance of the single-channel system relative to the two-channel system.

Yanick and Drucker (1976) performed a second series of experiments that were similar to the previous experiments of Yanick (1976b), except that the systems studied consisted of, in addition to the linear system and the two-channel system used previously, a second two-channel compression system. This second two-channel system differed from the first in that amplitude expansion (CR = 0.7) was included in the low-frequency channel below the compression threshold. The 6 listeners tested had mild to moderate sensorineural impairments (average loss of 25-60 dB at 0.5, 1, and 2 kHz) and sloping audiograms (12-25 dB/octave above 0.75 kHz). Key-word scores on Harvard sentences for the linear, compression, and compression-expansion systems were 46, 77, and 87% at S/N = 6 dB, and 22, 61, and 69% at S/N = 0 dB. These results thus appear to support the conjecture (for example, Villchur, 1973) that expansion can reduce the effects of background noise. The increase in intelligibility with expansion is difficult to understand, however, given the low S/N ratios used in these experiments. More specifically, since the level of the noise in the low-frequency channel was only 0 to 6 dB below the peak speech level, and the compression threshold in this channel was 30 dB below the peak speech level, the level of the noise in the low-frequency channel should rarely have fallen below the compression threshold and expansion should have had little effect. Also, the results obtained using the original two-channel system are not entirely consistent with those of the previous study. For example, a comparison of the scores obtained in the two studies with this system by subjects with sloping losses shows that scores were roughly 6 points lower in the current study despite the fact that the subjects in this study had less severe losses. Also, the scores obtained with expansion in the current study were, on the average, only 2 points greater than those obtained without expansion in the previous study.

Barford (1976) explored the hypothesis that multi-channel compression individually fitted to restore normal equal-loudness contours is superior to linear amplification. In a study involving five subjects with sensorineural losses, he compared each subject's own aid to a four-band linear system that was "optimally chosen and fitted" to each subject's loss, and to one-, two-, and three-channel compression systems designed to restore normal equal-loudness contours for tones. All subjects had bilateral losses, were hearing aid users, and had normal hearing at low frequencies (losses of 20

dB or less at 0.25 or 0.5 kHz) and significantly impaired hearing at high frequencies (losses greater than 50 dB above 2 kHz). The average audiogram for the five subjects showed losses of 22, 36, and 62 dB at 0.5, 1, and 2 kHz. Characteristics of the four experimental systems were based on equal-loudness contours that were related to normal equal-loudness contours via the normal low-frequency hearing of the impaired subjects. The optimal linear system (CH0) was determined (in an unexplained manner) on the basis of previous research (see discussion of Barfod, 1972, in Chapter II). According to our estimate, which attempts to take account of certain acoustic factors ignored by Barfod, the functional gain provided by this system was roughly flat above 750 Hz, but fell at roughly 18 dB/octave below 750 Hz. Each compression system had one linear low-frequency channel plus one (CH1), two (CH2), or three (CH3) high-frequency compression channels. The band limits, compression ratios, and relative gains of the channels were chosen to restore normal equal-loudness contours. Band limits (averaged over subjects) were roughly 0.9, 2.0, and 2.8 kHz for CH3, 0.9 and 2.0 kHz for CH2, and 1.2 kHz for CH1; and the CR values (averaged over subjects) in the various channels (in order of increasing frequency) were 1.0, 2.5, 4.8, and 5.6 for CH3, 1.0, 3.8, and 4.2 for CH2, and 1.0 and 4.7 for CH1. Since post-compression equalization was not used, gain adjustments could only be made between channels. The systems were compared both in quiet and with speech-spectrum noise added before processing at various S/N ratios, using nonsense CVC syllables spoken by a male. In all tests the input speech level was fixed at 65 dB SPL, and materials were presented using Beyer DT-48 headphones.

The relative performance of the five systems did not vary significantly with S/N ratio or subject (except for one subject whose results are excluded from the following averages). The scores, in terms of percent phonemes correct averaged over the quiet condition and S/N ratios of −5, 0, 5, 10, and 15 dB, were 54% (own aid), 65% (CH0), 46% (CH1), 46% (CH2), and 64% (CH3). Thus the three-channel compression system, which came closest to restoring normal equal-loudness contours, performed neither better nor worse than the "optimal" linear system; however, both of these systems were much superior to the one- and two-channel compression systems and significantly better than the subject's own aid.

Barford was the first investigator to study the effect of varying the number of compression bands systematically. However, the gain and compression ratio of each channel were not varied independently but rather co-varied to restore normal equal-loudness contours as well as possible for a given number of channels. Also, as the number of channels was reduced, the release times were increased (to avoid distortion at low-frequencies) and the attack times, which were constrained to equal the release times, were lengthened as well. It is not known whether his results would apply to systems in which these parameters were varied independently.

Barfod's results on compression are at variance with the more positive findings of Villchur (1973). Among the differences between the two studies that may account for this discrepancy are the following. First, Barfod's subjects on the average had sharply sloping losses; Villchur's subjects had more severe but gradually sloping losses. Second, Barfod's linear systems employed greater high-frequency emphasis than Villchur's. Third, Barfod's system was adjusted to restore normal equal-loudness contours for pure tones; Villchur's system included post-compression equalization and his procedures included individual modifications of system parameters. Fourth, Barfod's tests were performed at a fixed input level; Villchur's tests were performed at each subject's most comfortable level and 10 and 20 dB below this level. Fifth, Barfod's system had equal attack/release times that varied from 24 msec at low frequencies to 6 msec at high frequencies; Villchur's system had attack times of less than 1 msec and release times of roughly 20 msec in all channels.

D. CONCLUDING REMARKS

(1) Laboratory studies have demonstrated that compression-limiting preserves speech intelligibility at high levels better than simple peak clipping while providing the same protection from intense sounds. Recent clinical studies utilizing commercial hearing aids, however, have demonstrated no advantage for compression limiting. These negative results appear to be caused primarily by inappropriate compression characteristics in commercial aids and lack of individual fitting. Also, the aids compared have often differed in an uncontrolled manner with respect to a variety of electroacoustic properties other than the method used to limit intense sounds.

(2) Although the need for AVC in hearing aids has been recognized since 1950, there has been little serious research on this topic. There appear to be no technical obstacles to incorporating AVC in wearable hearing aids without compromising other essential features of the aids. Furthermore, the widespread use of AVC in sound recording applications suggests that, at least in certain highly-controlled circumstances, AVC can be used successfully without degrading quality or intelligibility. In order to determine the value of AVC for hearing aids, its effects on speech reception, comfort, and annoyance must be studied in controlled experiments with impaired listeners and in a wide variety of realistic environments. Testing AVC in realistic environments is particularly crucial because it is the function of AVC to control long-term levels, and constructing simulations of environments that preserve parameters relevant to the control of these levels is exceedingly difficult.

(3) Despite numerous studies of syllabic compression, it is still impossible to state whether syllabic compression is beneficial for persons with sensorineural losses.

Much of the research performed prior to 1970 was hampered by techni-

cal limitations that impeded the use of low compression ratios without employing instantaneous distortion devices, by poor design that introduced ancillary distortion and noise, and by the selection of inappropriate time constants. Also, almost all of the systems utilized single-channel compression in which neither the compression ratio nor the gain varied with frequency. In the only study prior to 1970 that employed multichannel compression (which, theoretically, should be superior to single-channel compression) and low compression ratios, the system suffered from considerable distortion and the compression ratio and gain were not varied over frequency.

During the last decade, it has become possible to overcome essentially all the technical limitations that faced earlier investigators. Advances in electronics have made it possible to build low-distortion compressors with compression ratios that are both low and variable. These compressors can now be easily combined to create multichannel compression systems that have low distortion, that are frequency dependent (in both compression ratio and gain) and that employ appropriate time constants. Unfortunately, however, the research conducted during the past decade with these more advanced systems has still not led to a clear picture of the benefits that can be achieved with syllabic compression. The results that have been obtained are in many cases inconsistent, and the reasons for the discrepancies are not yet thoroughly understood. Clearly, further research on syllabic compression is required before it can be recommended for clinical use.

REFERENCES

AMERICAN NATIONAL STANDARDS INSTITUTE, *Specification of hearing aid characteristics*. ANSI S3.22, New York: American National Standards Institute (1976).

ASPINALL, E., A master hearing aid. *J. British Institution of Radio Engineers*, 11, 45-50 (1951).

BARFOD, J., Investigations on the optimal corrective frequency response for high-tone hearing loss. The Acoustics Laboratory, Technical Univ. of Denmark, Rpt. No. 4 (1972).

BARFOD, J., Multichannel compression hearing aids: effects of recruitment on speech intelligibility. The Acoustics Laboratory, Technical Univ. of Denmark, Rpt. No. 11 (1976).

BEVAN, W. R., A professional mixer/compressor for remote or studio use. Preprint 916(B-3), 45th Convention of Audio Engineering Society (1973).

BIZAGUET, G., and VEIT, P., Use of advanced apparatus on deaf children: difficulties and results. *Rev. Laryng. Otol. Rhinol.*, 89, 448-456 (1968).

BLACKMER, D. E., A wide dynamic range noise-reduction system. *DB*, 64, 54-56 (1972).

BLEGVAD, B., Clinical evaluation of behind-the-ear hearing aids with compression amplification. *Scand. Audiol.*, 3, 57-60 (1974).

BLESSER, B. A., Audio dynamic range compression for minimum perceived distortion. *IEEE Trans. Aud. Electroacoust.*, AU-17, 22-32 (1969).

BLESSER, B. A., and KENT, A. R., Analysis of a feedback-controlled limiter using a logarithmic measuring scale. *IEEE Trans. Aud. Electroacoust.*, AU-16, 481-485 (1968).

BRINK, B., BORRILD, K., and OVING, K., Body-worn stereophonic compression aids used by profoundly hearing-impaired children. *Scand. Audiol.*, 4, 199-205 (1975).

BURCHFIELD, S. B., Perception of amplitude compressed speech by persons exhibiting loudness recruitment. Unpublished Ph.D. thesis, Michigan State Univ. (1971).

BURNETT, E. D., and BASSIN, M. A., Attack and release characteristics of compression hearing aids. NBSIR 76-1179, Washington: National Bureau of Standards (1976).

BURNETT, E. D., and SCHWEITZER, H. C., Attack and release times of automatic-gain-control hearing aids. *J. Acoust. Soc. Am.*, **62**, 784-786 (1977).

BURWEN, R. S., Design of a noise eliminator system. Preprint No. 838 (B-8), 41st Convention of the Audio Engineering Society (1971).

CARAWAY, B. J., Effects on speech intelligibility of dynamic range reduction with a compressor amplifier. Unpublished Ph.D. thesis, Northwestern Univ. (1964).

CARAWAY, B. J., and CARHART, R., Influence of compressor action on speech intelligibility. *J. Acoust. Soc. Am.*, **41**, 1414-1438 (1967).

CARTER, R. O., Theory of syllabic compandors. *Proc. IEEE*, **111**, 503-513 (1964).

DANIEL, D. B., The consonant amplifier-limiter. *Radio and Television News*, **57**, 49-51 (1957).

DAVIS, H., STEVENS, S. S., NICHOLS, R. H., HUDGINS, C. V., PETERSON, G. E., MARQUIS, R. J., and ROSS, D. A., *Hearing Aids–An Experimental Study of Design Objectives*. Cambridge: Harvard Univ. Press (1947).

DE VOS, A. W., Discussion following paper by Fleming and Rice. *Internat. Audiol.*, **8**, 523 (1969).

DOLBY, R. M., An audio noise reduction system. *J. Aud. Eng. Soc.*, **15** 383-388 (1967).

DUDLEY, H., and GRUENZ, O. O., Visible speech translators with external phosphors. *J. Acoust. Soc. Am.*, **18**, 62-73 (1946).

EDGARDH, B. H., The use of extreme limitation for the dynamic equalization of vowels and consonants in hearing aids. *Acta Otolaryng.*, **40**, 376-382 (1952).

EGAN, J. P., Articulation testing methods. *Laryngoscope*, **58**, 955-991 (1948).

FAIRBANKS, G., and MIRON, M. S., Effects of vocal effort upon the consonant-vowel ratio within the syllable. *J. Acoust. Soc. Am.*, **29**, 621-626 (1957).

FLEMING, D. B., and RICE, C. G., New circuit development concept in hearing aids. *Internat. Audiol.*, **8**, 517-523 (1969).

FOURNIER, J. E., *Audiometrie Vocale*, Librairie Maloine S. A., Paris (1951).

GREGORY, R. L., and DRYSDALE, A. E., Squeezing speech into the deaf ear. *Nature*, **264**, 748-751 (1976).

HATHAWAY, J. L., Automatic audio gain controls. *Audio Eng.*, **34**, September, 16-18 (1950).

HECKER, M. H. L., A study of the relationship between consonant-vowel ratios and speaker intelligibility. Unpublished Ph.D. thesis, Stanford Univ. (1974).

HELLWARTH, G. A., and JONES, G. D., Automatic Conditioning of Speech Signals. *IEEE Trans. Audio Electroacoust.*, **AU-16**, 169-179 (1967).

HIRSH, I. J., DAVIS, H., SILVERMAN, S. R., REYNOLDS, E. G., ELDERT, E., and BENSON, R. W., Development of materials for speech audiometry. *J. Speech Hearing Dis.*, **17**, 321-337 (1952).

HOOD, J. D., and POOLE, J. P., Tolerable limit of loudness: its clinical and physiological significance. *J. Acoust. Soc. Am.*, **40**, 47-53 (1966).

HOUSE, A., WILLIAMS, C., HECKER, M., and KRYTER, K., Articulation-testing methods: consonantal differentiation with a closed-response set. *J. Acoust. Soc. Am.*, **37**, 158-166 (1965).

HUDGINS, C. V., MARQUIS, R. J., NICHOLS, R. H., PETERSON, G. E., and ROSS, D. A., The comparative performance of an experimental hearing aid and two commercial instruments. *J. Acoust. Soc. Am.*, **20**, 241-258 (1948).

HUIZING, H. C., and REYNTJES, J. A., Recruitment and speech discrimination. *Laryngoscope*, **62**, 521-527 (1952).

IEEE, Recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust.*, **AU-17**, 225-247 (1969).

JOHANSSON, B., The hearing aid as a technical audiological problem. *Scand. Audiol., Suppl.* **3**, 55-76 (1973).

JOHANSSON, B., and LINDBLAD, A. V., The use of compression and frequency-transposition in hearing aids. In H. W. Erwertsen (Ed.), Electroacoustic Characteristics Relevant to Hearing Aids, *Scand. Audiol., Suppl.* **1**, 68-71 (1971).

KAISER, A., and BAUER, B. B., A new automatic level control for monophonic and stereophonic broadcasting. *IRE Trans. Audio*, **AU-10**, 171-173 (1962).

KREBS, D., Output limiting and related harmonic distortion. Paper presented at the 2nd Oticongress, Copenhagen (1972).

KRETSINGER, E. A., and YOUNG, N. B., The use of fast limiting to improve the intelligibility of speech in noise. *Speech Monogr.*, **27**, 63-69 (1960).

LUDVIGSEN, C., and NIELSEN, H. B., The effect of body-word stereophonic compression hearing aids on severely hearing-impaired adults. *Scand. Audiol.*, **4**, 191-197 (1975).

LYNN, G., and CARHART, R., Influence of attack and release in compression amplification on understanding of speech by hypoacusics. *J. Speech Hearing Dis.*, **28**, 124-140 (1963).

McCANDLESS, G. A., Hearing aids and loudness discomfort. Paper presented at the 3rd Oticongress, Copenhagen (1973).

NABELEK, I. V., On transient distortion in hearing aids with volume compression. *IEEE Trans. Aud. Electroacoust.*, **AU-21**, 279-285 (1973).

NABELEK, I. V., and ROBINETTE, L. N., Comparison of hearing aids with amplitude compression. *J. Acoust. Soc. Am.*, **57**, 519(A) (1975).

NOBLE, J. J., and BIRD, J. R., A dual-band audio limiter. *J. Aud. Eng. Soc.*, **17**, 678-684 (1969).

PARKER, C. D., The effects of the reduction of 'short-time fatigue' on speech intelligibility for 'perceptively' deafened individuals. Unpublished Ph.D. thesis, Univ. of Iowa (1953).

PESTALOZZA, G., Comparative study of the efficiency of various volume control devices of hearing aids. *Archives of Italian Otolaryng.*, **64/2**, 195-206 (1953).

PETERSON, G., and LEHISTE, I., Revised CNC lists for auditory tests. *J. Speech Hearing Res.*, **27**, 62-72 (1962).

POLIAKOFF, A., Hearing aid design. *Wireless World*, **56**, 274-276 (1950).

PORTMANN, M., and PORTMANN, C., *Clinical Audiometry*, Springfield: Charles C Thomas (1961).

RICHARDS, A. M., Loudness growth under masking: relation to true sensorineural impairment. *J. Speech Hearing Res.*, **16**, 597-607 (1973).

RINTLEMANN, W. F., Effects of amplitude compression upon speech perception: a review of research. *Scand. Audiol.*, **1**, 127-134 (1972).

ROBINSON, C. E., and HUNTINGTON, D. A., The intelligibility of speech processed by delayed long-time-averaged compression amplification. *J. Acoust. Soc. Am.*, **54**, 314(A) (1973).

RUHRBERG, W., and ESSER, G., Die wirkung der dynamik-kompression bei innenohrschwerhorightkeit und retrocochlearer horstorung. *HNO*, **21**, 247-255 (1973).

SALMON, R. D., Talker variation as related to intelligibility. Tech. Rpt. 31, Communication Sciences Laboratory, Univ. of Florida, Gainesville (1970).

SHORTER, D. E. L., MANSON, W. I., and STEBBINGS, D. W., The dynamic characteristics of limiters for sound programme circuits. *BBC Eng.*, Monogr. No. 70 (1967).

SILVERMAN, S. R., The implications for schools for the deaf of recent research on hearing aids. *Am. Ann. Deaf*, **94**, 325-339 (1949).

SILVERMAN, S. R., and HARRISON, C. E., The national research council group hearing aid project, a progress report. *Am. Ann. Deaf*, **96**, 420-431 (1951).

STEINBERG, J. C., and FRENCH, N. R., The portrayal of visible speech. *J. Acoust. Soc. Am.*, **18**, 4-16 (1946).

STEVENS, S. S., Power group transformations under glare, masking, and recruitment. *J. Acoust. Soc. Am.*, **39**, 725-735 (1966).

THOMAS, I. B., and SPARKS, D. W., Discrimination of filtered-clipped speech by hearing impaired subjects. *J. Acoust. Soc. Am.*, **49**, 1881-1887 (1971).

TILLMAN, T. W., CARHART, R., and WILBER, L. A., A test for speech discrimination composed of CNC monosyllabic words (N.U. auditory test no. 4). Tech. Doc. Rpt. No. SAM-TDR-62-135, USAF School of Aerospace Medicine, Aerospace Medical Division (AFSC), Brooks AFB, Tex (1963).

TILLMAN, T. W., and CARHART, R., An expanded test for speech discrimination utilizing CNC monosyllabic words (N.U. auditory test no. 6). Tech. Doc. Rpt. No. SAM-TDR-66-55, USAF School of Aerospace Medicine, Aerospace Medical Division (AFSC), Brooks AFB, Tex (1966).

TORICK, E. L., ALLEN, R. G., and BAUER, B. B., Automatic control of loudness level. *IEEE Trans. Broadcasting*, **BC-14**, 143-146 (1968).

TRINDER, E., An attempt to correct speech discrimination loss in cochlear deafness by graded instantaneous compression. *Sound*, **6**, 62-67 (1972).

VARGO, S. W., Intelligibility of amplitude compressed speech in quiet: normal and pathological hearing groups. *Audiology and Hearing Education*, February/March, 16-17 (1977).

VARGO, S. W., and CARHART, R., Amplitude compression, I. Speech intelligibility in quiet with normal and pathological hearing groups. *J. Acoust. Soc. Am.*, **53**, 327(A) (1972).

VIET, I., Practical possibilities in the design of hearing aids with automatic gain control. *J. Audiol. Tech.*, **12**, 186-200 (1973).

VILLCHUR, E., Signal processing to improve speech intelligibility in perceptive deafness. *J. Acoust. Soc. Am.*, **53**, 1646-1657 (1973).

WALLENFELS, H. G., *Hearing Aids on Prescription*, Springfield: Charles C Thomas (1967).

WILLIAMS, C. E., HECKER, M. H. L., STEVENS, K. N., and WOODS, B., Intelligibility test methods and procedures for the evaluation of speech communications systems. Rpt. ESD-TR-66-677, Decision Sciences Laboratory, Electronics Systems Division (AFSC) (1966).

YANICK, P., Jr., Improvement in speech discrimination with compression vs. linear amplification. *J. Aud. Res.*, 13, 333-338 (1973).

YANICK, P., Jr., Discrimination in the presence of competition with an AVC vs. a DRC hearing aid. *J. Am. Audiol. Soc.*, 1, 169-172 (1976a).

YANICK, P., Jr., Effects of signal processing on intelligibility of speech in noise for persons with sensorineural hearing loss. *J. Am. Audiol. Soc.*, 1, 229-238 (1976b).

YANICK, P., Jr., and DRUCKER, H., Signal processing to improve intelligibility in the presence of noise for persons with ski-slope hearing impairment. *IEEE Trans. Acoust. Speech Signal Processing*, ASSP-24, 507-512 (1976).

Chapter IV

PREVIOUS RESEARCH ON FREQUENCY LOWERING

A. PRELIMINARY REMARKS

In general, frequency lowering is addressed to listeners with sensorineural impairments who have extreme loss of hearing at high frequencies. That these listeners should have great difficulty in understanding speech signals is not surprising. For example, trained listeners with normal hearing typically achieve only 25-40% correct identification of isolated monosyllables when spectral components above 1 kHz are removed by filtering and suffer further degradations in performance when background noise is introduced (French and Steinberg, 1947; Pollack, 1948). Although in certain situations the loss of high-frequency acoustic cues can be compensated for by contextual information or lipreading, such strategies have obvious limitations. In addition, linear amplification or amplification combined with amplitude compression is only of limited use for such impairments. Among the possible features of such impairments that limit the usefulness of these processing schemes are the following. First, and most obvious, if the loss at high frequencies is essentially complete, it is impossible for the listener to detect the high-frequency sound components (except, perhaps, vibrotactilely). Second, even if the loss is not complete, the listener's resolution in amplitude, frequency, or time for these high-frequency components may be so poor that it is impossible to distinguish between different high-frequency sounds even when they are detected. Note also that if amplitude compression is used to compensate for a drastically reduced dynamic range at high frequencies, the listener's effective amplitude resolution and ability to make use of cues conveyed by relative intensity at high frequencies will be drastically reduced, independent of the intrinsic ability to resolve intensity. Finally, in addition to the inability to hear and resolve high-frequency components, such losses may entail various types of interfering abnormal sound distortions when the listener is presented with intense high-frequency stimulation.

Our review of previous work on frequency lowering is divided into six sections, according to the type of scheme employed: (B) Slow-Playback, (C) Time-Compressed Slow-Playback, (D) Frequency Shifting, (E) Vocod-

ing, (F) Zero-Crossing-Rate Division, and (G) Transposition. Concluding remarks are presented in Section H.

B. SLOW-PLAYBACK

When a sample of prerecorded sound is replayed at a rate slower than that used in recording, each spectral component is scaled lower in frequency by a multiplicative factor precisely equal to the slowdown factor. Because the time scale of the resultant waveform is dilated, the technique is not directly suitable for real-time applications. Nevertheless, it is important to understand the perception of speech processed by this technique because slow-playback is a component of many frequency-lowering schemes. In slow-played speech, the relation between the short-term spectral envelope and the fundamental frequency of voiced sounds is the same as for normal speech because all spectral components are lowered proportionally. The perception of slowed-played speech, however, is affected by changes in both spectral and temporal cues.

Investigation of the intelligibility of speech spectrally lowered by the slow-playback method for normals has focused on vowels (in an /h/-V-/d/ context), nonsense monosyllables, and phonetically balanced words, and has spanned the last 45 years (Fletcher, 1929; Ochiai and Izumitachi, 1955; Ochiai, Saito, and Sakai, 1955; Kurtzrock, 1956; Tiffany and Bennett, 1961; Daniloff, Shriner, and Zemlin, 1968). None of the studies has allowed the listeners used in the intelligibility studies more than about four hours of exposure to processed speech before testing. In all but one study, no feedback of correct response was provided during testing. Results obtained with normal listeners indicate that bandwidth reductions up to about 25% can be achieved with only slight loss in intelligibility, bandwidth reductions of 50% entail moderate intelligibility losses (30-40%), and bandwidth reductions of 66% or more entail severe intelligibility losses (70-80%). Several aspects of the deterioration in intelligibility appear significant. First, the deterioration appears to be less for voices of females and children than for males. This presumably reflects the fact that the spectral content of the speech of females and children is typically higher (both for formants and for fundamental frequency) than that of males. Second, for American English the perception of vowels deteriorates at least as rapidly as that of consonants with degree of frequency lowering. For Japanese, which has fewer vowels, consonants are affected more strongly than vowels. Third, American vowels having high first-formant frequencies and low second-formant frequencies are more severely affected than those with low first formants and high second formants. Fourth, one study (Tiffany and Bennett, 1961) indicates that repeated exposure to practice materials combined with feedback of correct responses can improve performance by a substantial amount over the course of four one-hour training sessions. This suggests that familiarization with the transformed speech cues is required before frequency lowered materials can be evaluated.

Bennett and Byers (1967) evaluated slow-playback speech with 15 listeners having high-frequency sensorineural hearing losses (average losses of 5, 7, 28, 51, and 66 dB at 0.25, 0.50, 1.0, 2.0, and 4.0 kHz). The intelligibility study used the Rhyme Test (Fairbanks, 1958) with female voice, but no training or feedback. Bandwidth reduction of up to 20% resulted in small improvements in scores (from 76 to 84%), but further lowering diminished performance substantially. The majority of impaired listeners reported understanding unprocessed male speech better than female speech and indicated that the slow-played speech generally sounded like male speech.

Except for certain small improvements obtained for modest bandwidth reductions with impaired listeners, bandwidth reductions have not been achieved by slow-playback without unacceptable loss of intelligibility (at least for untrained listeners). That this should be so has been noted by Pickett (1972): slow-playback produces sounds which are generally similar to the back vowels /u/, /ɔ/, and /o/ and to the low-frequency consonants /w/ and /r/. Untrained listeners should not be expected to interpret the modified sound patterns properly. Also, the time dilation inherent in slow-playback processing may, by itself, obliterate cues conveyed by the duration of transitions and vowels, and may complicate problems of short-term memory. Finally, the low intelligibility of slow-playback speech may reflect poorer resolution or greater spread of masking associated with low-frequency sounds. Unfortunately, the results of the tests reported thus far do not permit these effects to be evaluated separately.

C. TIME-COMPRESSED SLOW-PLAYBACK

The time dilation inherent in slow playback of prerecorded speech both renders the technique impractical for incorporation in real-time processing schemes and may obliterate cues keyed to temporal relations between speech sounds. A number of techniques for compressing speech in time are based on deleting segments periodically (Fairbanks, Everitt, and Jaeger, 1954), deleting successive pitch periods of voiced sounds (David and McDonald, 1956; Scott and Gerber, 1972), and deleting segments in adherence to phonological rules (Toong, 1974). Although some of these techniques undoubtedly introduce sampling noise and other distortions, they indicate that much speech redundancy can be eliminated on a temporal basis without reducing intelligibility. Up to roughly 80% of the waveform can be deleted without substantial loss of intelligibility if the deletions are short enough (Miller and Licklider, 1950) or are accomplished pitch synchronously and padded with repetitions or linear interpolations of the retained pitch periods (David and McDonald, 1956). When the retained segments are abutted in time, resulting in a reduction in overall duration, the loss in intelligibility is small for time-compression factors on the order of 2-4 (Fairbanks, Everitt, and Jaeger, 1954). However, the waveform discontinuities associated with the abutting of nonadjacent segments are generally very objectionable and may reduce intel-

ligibility unless some form of smoothing is employed (Bennett and Linvill, 1975). Furthermore, the comprehensibility of time-compressed speech may not be well predicted by the intelligibility of isolated time-compressed words (Foulke and Sticht, 1969).

Time-compressed speech used in frequency-lowering studies is processed by the slow-playback technique discussed above with compression and slowdown factors generally chosen to be equal. Under these conditions, the short-term spectrum is lowered and the long-term duration of speech is preserved. Speech processed in this way retains many temporal and rhythmic cues, and thus should be more easily understood than slow-played speech. The effect of time-compressed slow-playback processing on fundamental voice frequency depends crucially on the length of the discard interval and the discard rate. If the discard rate is slow relative to the fundamental frequency, with a correspondingly long discard interval, then the processing lowers the fundamental frequency. If the rate is precisely equal to the fundamental frequency, and the interval a fraction of the fundamental period (that is, the processing is pitch-synchronous), no alteration in fundamental frequency occurs (Stover, 1967). If, however, the rate is merely comparable to the fundamental frequency, but not synchronous, the processing may distort the periodicity of the waveform severely. The discard interval has been varied systematically in studies involving normal listeners. Good results have generally been obtained for intervals of 20-60 msec. Studies with impaired listeners have generally not explored this parameter, but have used intervals much longer than the fundamental period. No pitch-synchronous lowering studies appear to have been reported either for normal or impaired listeners.

The time-compressed slow-playback approach to frequency lowering has been evaluated using normal adult listeners with three hours of training or less for phonetically balanced words (Fairbanks and Kodman, 1957; Klumpp and Webster, 1961), vowels in an /h/-V-/d/ context (Zemlin, 1966; Daniloff, Shriner, and Zemlin, 1968), and Japanese monosyllables (Nagafuchi, 1976). A few studies (for example, Shriner, Beasley, and Zemlin, 1969; and Nagafuchi, 1976) have tested young school-age children. The results indicate that time-compression improves the intelligibility of slow-played speech only slightly, if at all; that larger reductions are possible for female speech than male speech, as in the lowered-dilated case; and that perception of vowels is degraded to an equal or greater extent than consonants. As in time-dilated frequency-lowering, bandwidth compression by factors of 2 and 3 were found to decrease intelligibility by about 50% and 80% respectively.

Knorr (1976) has proposed that processing to achieve frequency lowering of unvoiced sounds should be incorporated in hearing aids for persons with extreme high-frequency losses. In this scheme, voiced sounds are merely amplified, while unvoiced sounds are lowered in frequency by slow-playback (reduction factor of 4) combined with periodic discarding. The relative level of the processed and unprocessed signals is adjustable.

Although pilot versions of the proposed aid have been built, no evaluation of its performance relative to linear amplification have been presented.

Oeken (1963) evaluated the time-compressed slow-playback technique using 40 listeners with high-frequency sensorineural deafness (at least 40 dB loss between 1200 and 1500 Hz). In the absence of training, lowering the spectrum ½ to 1 octave was found to decrease intelligibility scores for numbers, words and phrases, and to decrease the comprehension of short stories. When four young patients were trained intensively (by unspecified techniques) on the processed speech, word intelligibility increased by 30% or more. However, corresponding training with unprocessed speech achieved similar results and normal speech always proved more intelligible than spectrally-lowered speech.

Haspiel (1969), in a preliminary study of a time-compressed slow-playback processing scheme designed by Searle (1969), determined that hard-of-hearing adults with mild to moderate losses could be trained to achieve significant improvements in word-recognition scores in 12 hours of listening time, even for the relatively high lowering factors of 2 and 2.5. Redden (1973) has conducted word recognition tests (W-22 lists) on large groups of untrained college students with normal hearing using materials processed by the Searle technique. Average scores range from 83% for materials processed with no frequency lowering to 57% for a lowering factor of 1.5, and 2% for a lowering factor of 3. Analysis of word-recognition errors in terms of phonemic confusions showed that each lowering factor induced a distinct pattern of errors; however, this finding was undoubtedly due in part to the absence of training and feedback and the use of separate groups of listeners for each reduction factor.

Zemlin (1966) reported discouraging results from a preliminary evaluation of the time-compressed slow-playback technique using grade-school children. With normal hearing children, W-22 test scores were 67% for unprocessed materials, and 50, 26, and 12% for lowering factors of 1.4, 1.7, and 2.0, respectively. With hearing-impaired children, vowel recognition scores (in /h/-V-/d/ context) were 55% for unprocessed materials and 38% for a lowering factor of 2.0. Zemlin interpreted these results, which generally reflect tests conducted on untrained listeners, as indicating that frequency lowering beyond a certain amount so transcends a listener's previous language experience that intelligibility is reduced even for listeners with hearing handicaps who would be expected to benefit from the bandwidth reduction.

Beasley, Mosher, and Orchik (1976) studied the ability of hearing impaired children to learn to identify words processed by an electronic system (Lee, 1972) to achieve a bandwidth reduction of 35% through slow-playback and time compression. Eighteen children (6-10 years of age) with congenital sensorineural hearing losses (average losses of 90, 101, and 104 dB at 0.5, 1.0, and 2.0 kHz) were paired and divided into two groups. One group (the control group) was trained on linearly amplified speech, the other on frequency lowered materials. Each child was trained

for 20 minutes/day on 15 consecutive school days. A comparison of identification scores before and after training indicated that the group trained on the frequency-lowered materials improved more than the control group on both the frequency-lowered and linearly amplified materials. However, final scores for this group were roughly the same for both types of processing, and were similar to those of the control group for the linearly amplified materials. Thus, although the authors concluded that frequency-lowering has potential for application to the education and auditory training of hearing impaired children, their data suggest that the advantage of frequency lowering over linear amplification is marginal at best.

Mazor, Simon, Scheinberg, and Levitt (1977) studied the intelligibility of PB-50 words processed by the Lee (1972) system for adults. Eight listeners with sensorineural losses (average losses of 25, 30, 45, 53, and 63 dB at 0.25, 0.50, 1.0, 2.0, and 4.0 kHz), and four normal-hearing adults with losses simulated by 1.5 kHz low-pass filtering, were tested at spectral reductions of 20, 33 and 55% for both male and female speech. Training consisted of listening to 40 minutes of processed materials spoken by the voices used in the test. In all test conditions but one—20% reduction, female voice, impaired listeners—spectral lowering reduced intelligibility relative to unprocessed materials and the reduction increased as the lowering factor increased. For the exceptional case, a small average improvement in scores (3 percentage points) was found, although not all impaired listeners showed increased scores. Also, the impaired listeners found processed female speech more intelligible than processed male speech for each reduction factor, while the normal listeners found processed male speech more intelligible for the 20% reduction factor.

Schreiner (1977) has suggested that persons with profound high-frequency hearing loss might benefit from a combination of frequency-lowering, amplitude compression, and linear amplification. In his system, frequency lowering is achieved by pitch-synchronous time-compressed slow-playback processing in which 3 out of 4 successive pitch periods are discarded and the remaining period dilated by a factor of 4. Broadband amplitude compression and high-frequency emphasis are included to compensate for the reductions in sensitivity and dynamic range characteristic of profound high-frequency loss. Schreiner noted that speech processed by this system was highly unnatural in character, perhaps partially due to the lowering of fundamental frequency by a factor of 4, and would require substantial training in order to evaluate recognition performance for processed materials. Also, speech sounds of very short duration were found to be transformed non-uniquely by the processing, perhaps due to the long discard intervals—typically 30 msec for male speech, while longer duration sounds were transformed in a more perceptually constant fashion. In order to avoid the necessity of extended training periods, a preliminary evaluation of this system utilized a speech sound discrimination test in which listeners were merely required to de-

termine the odd member of a triplet without identifying the speech sounds. Eight listeners with normal hearing, who heard materials filtered through a 50-800 Hz passband to simulate a hearing loss, were tested on discrimination of 20 vowels and diphthongs (in /h/-V-/d/ context) and 23 consonants (in C-/ai/ context). Average scores for vowels (92%) were substantially better than for consonants (75%). Furthermore, although most vowel confusions occurred within the class of tense-back vowels, nearly all consonants were confused with one another. Schreiner argued that these results are sufficiently encouraging to warrant further development of the system and evaluation as a training or reception aid for impaired listeners.

The results on spectral lowering achieved by slow-playback with time compression are very similar to those achieved without time compression. This is perhaps not surprising since the test materials consisted exclusively of short utterances for which time-compression may be relatively unimportant. Also, independent of whether time compression is used, the only positive results reported have been for female voice and for small amounts of frequency lowering, and in these cases the average improvements have been only very modest.

D. FREQUENCY SHIFTING

Frequency lowering can also be achieved in real time by heterodyne techniques, which use amplitude modulation to shift all spectral components in a given frequency band downward by a fixed displacement. This shifting involves no bandwidth reduction per se, and thus generally produces a signal in which high-frequency components and low-frequency components are overlaid or "aliased." When aliasing occurs, the direction of frequency transitions may be reversed. Aliasing can be eliminated, of course, by removing the low-frequency components of the signals by filtering before modulation. Heterodyne systems which shift only high-frequency components (for example, 3-6 kHz) are generally referred to as transposers and will be discussed in Section G. below. This section is confined to systems which shift the entire spectrum, except for those low-frequency components eliminated by filtering (for example, 0-800 Hz).

Although speech lowered by heterodyning retains the temporal and rhythmic intensive patterns of unprocessed speech, the harmonic relationships characteristic of voiced sounds are greatly altered. Thus, voice pitch is strongly modified and may even become ambiguous. Further, the prefiltering used to prevent aliasing obliterates the low-frequency components of speech. For example, in order to lower all frequencies less than 5 kHz by 20% without aliasing, all components below 1 kHz must be eliminated. Such filtering by itself would degrade intelligibility.

Fletcher (1953) studied the effects of frequency shifts on intelligibility for listeners with normal hearing. His data indicate that both positive and negative shifts reduce intelligibility, but that shifts toward lower frequen-

cies are more deleterious than equivalent shifts toward higher frequencies. Shifting the spectrum lower by 300 Hz reduces intelligibility by roughly 35%—about the same extent as slow-playback at ⅔ normal rate. Note, however, that a 3000 Hz spectral component would be lowered only 300 Hz by heterodyning whereas it would be lowered by 1000 Hz by slow-playback.

Campbell, Shultheis, and Barton (1976) have suggested that when residual hearing exists only at very low frequencies, frequency shifting may be beneficial and yield performance superior to slow-playback. In tests of normal listeners with losses simulated by 500-Hz lowpass filtering, significantly better scores were achieved when speech materials were shifted lower by 500 Hz (45%) than for unprocessed speech. No advantage was found for time-compressed slow-playback processing (including frequency-lowering factors of 1.5 and 2.0) for these listening conditions.

Raymond and Proud (1962) have evaluated the intelligibility of frequency-shifted speech using 16 listeners with predominantly high-frequency impairments (average losses of 16, 14, 18, 39, 52, and 64 dB at 0.25, 0.50, 1.0, 2.0, and 4.0 kHz). The subjects were trained a total of 16 hours, spread over two months, by an experienced remedial reading teacher. Training consisted of listening to simple stories and word tests, and included both diotic listening to shifted speech and dichotic listening with shifted speech presented to one ear and normal speech to the other. After training, word-intelligibility tests were administered at equal sound levels: with a female speaker and no frequency shift, with a male speaker and a frequency shift of 400 Hz, and with a female speaker and a frequency shift of 750 Hz. Tests included both diotic and dichotic listening. Results of the tests indicated substantial variability among subjects, particularly in dichotic listening. All listeners seemed to improve performance with training and all performed better with a female voice shifted by 750 Hz than with a male voice shifted by 400 Hz. However, none achieved significantly better scores for shifted speech after training, no matter how presented, than for unshifted male speech before training.

Biondi and Biondi (1968) have advocated the use of sampling to achieve frequency-shifting in aids for listeners with very large hearing losses. Their technique depends on aliasing to achieve frequency lowering. Although they claim that sampled speech is readily intelligible to normal listeners, this statement may apply only to sampling frequencies that are much higher than those used with impaired listeners and for which the aliasing effects are small. Furthermore, they have reported that some moderately impaired listeners who use simple amplifying aids were so negatively impressed by the sampling technique that they refused to cooperate in its evaluation. On the other hand, five severely impaired (average losses of 60, 70, and 80 dB at 0.25, 0.50, and 1.0 kHz) young adult male subjects were reported to be capable of recognizing extensive vocabularies of words both in isolation and in sentence context after 30-60 hours of training. One subject (whose hearing loss was reported to exceed

120 dB at all frequencies!) reportedly became capable of recognizing any spoken word but could only be trained to remember about 30 words at any one time.

Ling (1972) has evaluated the pairwise discriminability of CV sounds processed by the Biondi and Biondi system using 12 severely impaired (average losses of 65, 80, and 90 dB at 0.25, 0.50, and 1.0 kHz) children. Discrimination scores showed no advantage for the processed sounds; in fact, linearly-amplified sounds were consistently more discriminable (although the differences were small).

Biondi and Biondi (1973) have recently reported further data on the performance of their system. With training, listeners with normal hearing have been found capable of recognizing, with 75% or greater accuracy, words selected from a fixed vocabulary of 20 items for a variety of sampling frequencies. Biondi and Biondi report that with extensive training impaired listeners can make use of the sampling technique to engage in conversation to a limited extent. Taken as a whole, the results obtained by Biondi and Biondi indicate that an extensive period of training is required before the severely impaired can begin to make use of the cues provided by the sampling technique. Further, the best results achieved indicate that only very limited vocabularies of sounds can be acquired.

In general, frequency-shifting techniques do not appear to offer much prospect of improving speech perception for the impaired. No encouraging results have been reported when small frequency shifts are used, in contrast to the positive results obtained with small degrees of slow-playback. The only positive results reported for frequency-shifting are related to the use of sampling techniques, together with extensive training, in the rehabilitation of the profoundly deaf. These results are very restricted: normal and moderately impaired listeners appear to find the sounds produced by the sampling technique very difficult to interpret as speech.

E. VOCODING

Channel Vocoders, first introduced by Dudley (1939), are used to achieve bandwidth reduction in certain voice communication systems. In such applications, speech is analyzed by a bank of contiguous bandpass filters whose output envelopes are detected and low-pass filtered for transmission. The received signals are used to control the amplitude of corresponding channels. These vocoders also transmit a signal descriptive of the vocal sound source (hiss or buzz) and the fundamental frequency of voiced sounds. Typically, transmission bandwidth reductions of a factor of 10 can be achieved with good intelligibility, although the bandwidth of the reconstructed signal is generally the same as that of the speech signal before encoding. The Channel Vocoder can be used to achieve frequency lowering by using a set of synthesis filters corresponding to lower frequencies than the associated analysis filters.

Channel Vocoders should have significant potential for lowering the spectra of speech signals for the hearing impaired. They are capable of operating in real time (with perhaps a small delay between analysis and synthesis) and are known to be capable of reproducing speech with good intelligibility. Further, as Risberg (1969) has pointed out, they provide a very flexible solution to the problem of frequency lowering since both the analysis and synthesis filters can be selected for a given application, and the excitation function used for synthesis can be specified independently. Thus, for example, it is possible to obtain transformations in which the spectrum of speech is lowered nonuniformly or in which the fundamental frequency of voiced sounds is manipulated independently of the spectral envelope. Although channel vocoders are not typically used to produce synthesized speech of lowered bandwidth for listeners with normal hearing, a related device, the harmonic compressor (Schroeder, Logan, and Prestigiacomo, 1972), has been used by the American Federation of the Blind to lower the spectrum of materials played back at twice the normal rate (Breuel and Levens, 1969). In this case, the resulting spectrum approximates that of normal speech but the duration of speech elements is halved.

Denes (1967) reported on tests of speech that was spectrally lowered by a factor of roughly three by an 11-channel vocoder. The scheme lowered voice pitch by a factor of three and compressed the spectrum nonuniformly such that high frequencies were lowered proportionately more than low frequencies. Denes evaluated the transformation on ten listeners with normal hearing, half of whom were allowed to listen to their own voices processed by the vocoder while the other half were exposed only to prerecorded processed isolated words taken from a list of 150 words. Both groups participated in 16 training sessions (of 20 minutes duration each) and were tested at the end of each session on 50 words. Although Denes reported no clear cut differences in final scores between the two groups, the spectrally-lowered speech was found to be learnable, with recognition scores improving from 40% to 70% after the 5.3 hours of training. Furthermore, he found that almost three times as many errors concerned place of articulation than manner of articulation. He attributed this result to rough quantization of the second formant region by the analyzing filters.

Takefuta and Swigart (1968) evaluated a channel vocoder in which the output consisted of 22 sinewaves whose amplitudes were modulated by envelope waveforms derived from one-third octave analysis filters. Word recognition tests using materials spoken by a female voice were presented both before and after a 15-minute training session. Four groups of twenty normal-hearing listeners were tested with materials filtered to a 50-9500 Hz bandwidth and lowered by factors of 1.0, 1.4, 2.0, or 2.5. Scores improved substantially after training for all conditions except the 2.5 reduction ratio, for which only minor improvement occurred. After training, scores averaged 70% for a ratio of 1.0 and 56% for a ratio of 2.0. An analy-

sis of the errors made indicated that consonants were perceived more correctly than vowels after processing. Although the testing focused on isolated words, the authors stressed that the coding scheme would degrade intonation patterns in running speech.

Minami and Katsurayama (1968) have reported on a channel vocoder in which the output consisted of six amplitude-modulated tones with frequencies below 2000 Hz. In this system no distinction is made between voiced and unvoiced sounds either in terms of analysis or synthesis. No performance results have thus far been reported. Pimonow (1963, 1968) has constructed a similar seven-channel system using tones below 300 Hz. He reports that three "deaf" 12-year-olds were able to learn a vocabulary of roughly 30 words within a period of three weeks' training. Lafon (1967) has reported development of a three-channel vocoder for the severely deaf. In Lafon's vocoder sounds below 1000 Hz are amplified conventionally to provide information about rhythm and melody. Two frequency bands, 1500-3000 Hz and 5000-7000 Hz, are analyzed by separate filters, whose output envelopes are used to modulate (unspecified) low-frequency sounds which are then added to the conventionally amplified channel. No details on the effectiveness of this scheme have been provided.

Ling and Druz (1967) constructed a six-channel vocoder system for use by deaf children with no hearing above 2 kHz. In this system, sounds in the 2-3 kHz range are analyzed by five bandpass filters whose output envelopes are used to modulate the amplitudes of five sinewaves in the 750-1000 Hz range. These processed signals are added to linearly amplified components in the 70-700 Hz range. All other sounds are eliminated by filtering. The intelligibility of speech processed by this system was compared to that of speech amplified linearly by a speech training aid with a smooth frequency response from 60-6000 Hz. Eight hard-of-hearing children (ISO HL's of 70, 81, 86 dB at 125, 250, and 500 Hz) were trained on either vocoded or linearly amplified speech for 22 days in periods of 20 minutes per day. Tests used materials prerecorded by a female speaker and were administered both before and after training using materials processed both by amplification and by vocoding. Five tests were administered: discrimination between back, mid, and front vowels in familiar words; discrimination between classes of consonants preceding a long vowel; discrimination within classes of consonants preceded by a short vowel; discrimination between intonation patterns; and phonetically balanced word lists. The results indicated that the children made significant improvements in speech discrimination, but that the improvements made by those trained with conventional amplification were somewhat larger than by those trained with the vocoder. However, final scores for materials processed in the same manner as the training materials were about the same for amplification and vocoding. In interpreting these results, it is important to note that the children who participated in the evaluation had very large losses in those frequency regions used to convey the recoded information.

Ling and Doehring (1969) constructed and evaluated a 10-channel vocoder similar to Piminow's. Sounds from 1.0 to 4.0 kHz were analyzed by 10 logarithmically spaced filters and used to modulate the amplitude of 10 tones spaced at 100-Hz intervals from 100 Hz to 1000 Hz. Twenty-four children aged seven to 11 years, known to be profoundly deaf from birth or early infancy, served as subjects. These children were divided into four groups (equated as closely as possible in age, hearing level, teachers' ratings, and pretest scores with linear amplification) and trained on consonant discrimination using an automated programmed instruction system. One group was trained with diotic linearly amplified sound, one with diotic vocoded sound, one with a dichotic presentation of amplified and vocoded sound, and one, which served as a control, with no auditory cues. Training occupied 15 minutes per day and proceeded until a limit of learning was attained. Pre- and post-training speech discrimination tests, using words spoken by the same female voice who prepared the training materials, were administered both with amplification and with vocoding. The results of these tests indicated that all groups learned to discriminate the training materials very well, but post-training scores were only slightly different from pre-training scores. All groups achieved better post-training scores on linearly amplified materials than on vocoded materials. Furthermore, only the group trained diotically on vocoded materials improved in discriminating vocoded speech. Thus, although the results do not indicate that frequency-lowered speech is more discriminable than linearly amplified speech, they do demonstrate that deaf children can learn to discriminate vocoded sounds.

The utility of the Ling and Doehring vocoder as an aid to articulation training was evaluated by Ling and Maretic (1971) on 18 congenitally impaired (average losses of 69 and 92 dB at 250 and 4000 Hz) children aged seven to 11 years. The children were trained to imitate pronunciation in short (15 minute), individual training sessions over a period of 40 days. During training the subjects were grouped and received either a diotic presentation of conventionally amplified sounds or a dichotic presentation of conventionally amplified sound and vocoded sound. Training and test materials were identical and consisted of CV nonsense syllables. All subjects were tested on each amplification condition both before and after training. All groups improved imitation skills as a result of training. Furthermore, the improvement occurred for all amplification conditions, not only the one used in training. Independent of training, no final test scores for vocoded materials were significantly superior to those for conventional amplification, suggesting that the vocoded signals were not used to supplement the information in the conventionally amplified signals.

An alternate vocoding strategy that has been proposed for voice communication systems is based upon estimation of the resonant frequencies of the vocal tract. In these formant vocoders (for example, Flanagan and House, 1956), the transmitted signals describe the time course of these resonances and are used at the receiver to control a formant synthesizer.

Formant vocoding has been used as the basis of a frequency-lowering scheme by Reeder (Reeder, 1975; Reeder, Strong, and Palmer, 1975). In this system, estimates of the first three formant frequencies are derived using a linear-prediction-based technique described by Strong and Palmer (1974). These estimates are used to modulate the frequencies of three sinewaves in the 100-1000 Hz range. During unvoiced speech segments, noise modulation was used at the transposed formant frequencies. Two listeners with normal hearing who trained on materials processed in this fashion for a period of 17-20 hours were able to acquire a vocabulary consisting of 50 common words. In discrimination tests (Voiers et al., 1973), these listeners achieved scores of 89-92% for consonantal features such as nasality, sustention, sibilation, graveness, and compactness and scores of 96% for voicing. Listeners with high-frequency losses were less able to discriminate the processed materials and achieved scores of only 81% for voicing and 58-65% for the other features. Reeder, Strong, and Palmer have suggested that the performance of this system could be improved by supplementing the three sinewaves with either low-frequency speech components or with a fourth sinewave that encodes the fundamental frequency of voiced sounds.

Stewart, Strong, and Palmer (1976) have evaluated a variant of this technique in which the frequency-lowered signal consists of four sinusoidal components synthesized in pitch synchrony with the original signal. The ability of untrained listeners with normal hearing to discriminate monosyllabic words was tested using a "same-different" version of the diagnostic rhyme test. The results indicate that the perception of voicing is only slightly degraded by frequency-lowering, but the perception of nasality, sibilation, and graveness were adversely affected. The authors suggest that the perception of sustention could be improved by achieving better time resolution in the signal processing.

In view of the obvious flexibility of vocoders, the lack of demonstrated positive results is disappointing but not too surprising. In Denes's system, for example, the sound quality may not have been significantly better than that found in slow-played time-compressed speech because the fundamental frequency was lowered by the same factor as the bandwidth reduction. Similarly, many of the vocoders proposed or evaluated fail to distinguish between voiced and unvoiced sounds and hence may produce sounds that are hard to interpret. Also, some of the vocoders were evaluated by listeners who had little residual hearing capacity, and probably very little ability to make fine discriminations among the coded signals. Despite these problems, several studies of the vocoder technique have found that discrimination and recognition of the frequency-lowered materials improved significantly with training.

F. ZERO-CROSSING-RATE DIVISION

A number of approaches to frequency lowering are based on attempts to

reduce the zero-crossing rate of the speech waveform. Schroeder, Flanagan, and Lundry (1967) have employed zero-crossing-rate division as a component of "analytic-signal rooting" to process speech for transmission over narrow-band channels. In this scheme, speech is filtered into four contiguous passbands, each roughly as wide as the spacing between formant frequencies, and the resulting signals are processed to achieve a band reduction by a factor of two. When these signals are reprocessed at the receiver, the resulting signal preserves the formant structure and fundamental-frequency information of the original speech and the overall transmission quality is relatively high. However, certain other signals, such as broadband noise, appear to be markedly distorted during transmission. There do not appear to have been any attempts to apply the analytic-signal-rooting technique directly to the problem of frequency lowering for impaired listeners (most likely because the signal processing required is quite complex), but a number of investigators have attempted to approximate the relevant effects. In these schemes bands of speech are extracted by filtering and the filter outputs converted to lower frequency sounds having reduced zero-crossing rates. In some of these schemes the low-frequency signals are formed by generating one pulse for each R zero-crossings of the filtered signal. In other schemes the low frequency sounds are derived by modulating the instantaneous frequency of a sinusoid with a signal derived from the instantaneous zero-crossing rate of the original signal. In some cases the amplitude of the low-frequency sounds is modulated by the envelope of the original signal, in others no such amplitude information is provided.

Guttman and van Bergeijk (1959) evaluated a three-channel zero-crossing-rate division system capable of lowering the speech spectrum by a factor of two on a single hearing-impaired subject (70 dB loss up to 1 kHz, no hearing beyond 1 kHz). The frequency-lowering system used consisted of the transmitting portion of the VOCBANC developed by Bogert (1956). Although Bogert reported that consonant articulation for speech processed by VOCBANC, low-pass filtered (to 1700 Hz), and translated to normal frequencies by the VOCBANC receiver, was somewhat superior to speech low-pass filtered to the same bandwidth, Schroeder (1966) has characterized the quality of the resultant speech as "rather poor." When Guttman and van Bergeijk used the VOCBANC for frequency lowering, they merely amplified the low-pass filtered signals and presented them to the listener. Although they reported that at comparably high-intensity levels the intelligibility of processed speech was approximately equal to that of natural speech, after several months of two-hours-per-week training, only very limited learning had occurred. The study was terminated because the cost in terms of time and facilities was prohibitive.

Guttman and Nelson (1968) tested a zero-crossing-rate-division system intended to recode fricative sounds in which the processing was applied only to high-frequency components. In this system a "voicing-detector" inhibits processing of voiced speech segments. During unvoiced seg-

ments components in the 100-2000 Hz region are extracted by bandpass filtering and added to processed components derived from the 6-15 kHz region. These processed components are zero-crossing synchronous pulses with R in the range 1-64. Spectrograms of the processed sounds corresponding to /s/ and /ʃ/ indicate that for R = 64, considerable power is introduced below 500 Hz, producing sound patterns which should be highly discriminable. Speech processed by this system and then low-pass filtered (cut-off frequencies in the range 0.6-1.2 kHz) to simulate hearing loss was evaluated by four normal listeners. Although the intelligibility of the processed speech was higher when the test materials consisted of small sets of words, there was little difference in intelligibility between the processed and control speech when the material consisted of larger word vocabularies or sentences. Guttman and Nelson interpreted these results to indicate that zero-crossing-rate division is unlikely to be a useful aid to speech reception for those suffering high-frequency losses.

Guttman, Levitt, and Bellefleur (1970) evaluated the system developed by Guttman and Nelson as an aid to the articulation training of six severely hard-of-hearing (average losses of 79, 85, 94, 105 dB at 125, 250, 500, and 1000 Hz) children (aged 10-13). These children were sufficiently impaired that they achieved a score of only 5% on amplified phonetically-balanced words (however, they could achieve a score of 66% by lipreading). Training occurred in 70 20-minute sessions on successive school days and consisted of /s/-/ʃ/ discrimination, articulatory drill, and practice on imitating fricative sounds. Both before and after training the subjects were asked to pronounce a set of words presented in random order. Ratings of the quality of the pronunciation of these words by naive listeners indicated that 51% of the pronunciation of words without /s/ or /ʃ/ and 81% of those with /s/ or /ʃ/ were rated better after training. However a control group trained on linearly amplified speech achieved somewhat larger improvements (63 and 93%). Quality ratings by a trained evaluator indicated that the Guttman-Nelson system may have been somewhat more effective in training /ʃ/ but not /s/. However, students who used the Guttman-Nelson system improved less on the articulation of other consonants than those in the control group. Ratings by the speech trainer during training indicated that, although the two groups could discriminate /s/ from /ʃ/ equally well, students using the Guttman-Nelson device appeared to achieve superior articulation of both /s/ and /ʃ/. Thus, these results indicate that processing by the Guttman-Nelson system, together with extensive training, resulted in improved articulation for selected phonemes. Equivalent amounts of training using linear amplification improved articulation somewhat less but over a larger class of phonemes. It is interesting to note that Bouillon (1968) found a similar preference among teachers of the deaf for speech processing systems that proved to have no advantage over other systems in terms of discrimination.

Thomas and Flavin (1970) tested a zero-crossing-rate division system in which the processing was applied to sounds in the 1-2 kHz octave band.

In their system, R = 2 and the envelope of the filtered signal was used to modulate the amplitude of the low-rate pulses. The Rhyme Test (Fairbanks, 1958) was administered to 11 normal-hearing subjects (who received neither special training nor trial-by-trial feedback) to evaluate the processed speech. Low-pass filtering (1.25 kHz) was used to simulate a hearing loss. The total exposure and testing time was about five hours. Average scores improved from 43% correct on the first test list to 77% correct on the last ten lists. Further, the confusions reported for a typical subject indicate that improvements in identification of both place and manner of articulation occurred during the course of testing.

Morgan and Savidge (1975) have developed a bandwidth compression-expansion scheme based on instantaneous frequency division-multiplication of an amplitude-modulated high-frequency carrier. Although this scheme has not been applied to the hearing impaired, bandwidth reduction factors of 5-10 are claimed to be in use for a compression-expansion system capable of transmitting voice bandwidth.

Done and Kirlin (1975) have proposed a similar technique in which the lowered signal consists of a tone that is amplitude-modulated by the envelope of the original signal and that is frequency-modulated by a signal derived from the instantaneous zero-crossing rate of the original signal. In this system the relation between zero-crossing rates can be easily adjusted and need not be linear. Although the relation between the spectra of the original and processed signals has not been reported, measurements of instantaneous zero-crossing rates suggest that the scheme successfully lowers the frequencies of the formant peaks of vowels. The intense fricatives /ʃ/ and /s/ are lowered more accurately than /f/ and /θ/. Preliminary listening tests indicate that word recognition is significantly degraded by the processing. Without training, listeners achieved scores of 98% for unprocessed materials, 82% for processed materials that are not lowered in frequency, and 8% for materials lowered by a factor of about 3.

Zero-crossing-rate division, when used to produce surrogate low-frequency components of fricative sounds, appears to have potential applications in limited speech training situations. No positive results have been reported for any zero-crossing-rate-division technique when used to improve intelligibility for the impaired.

G. TRANSPOSITION

The only frequency-lowering technique currently available in commercial hearing aids (Nielsen, 1972), and perhaps the most widely evaluated frequency-lowering scheme, involves transposition of the high-frequency components of sound to lower frequencies and addition of the transposed signals and the unprocessed low-frequency signals. The high-frequency components are derived by linear filtering and the transposition is achieved by amplitude modulation or nonlinear distortion. A transposition device has been patented by Hopner and Andrews (1962) to improve the

"consonant response in the transmission of speech signals" over narrow-band (300-3400 Hz) communication channels. No evaluation of the effectiveness of this device appears to be available in the public domain.

Johansson (1961) designed one of the first transposers intended for use by the hard of hearing. This transposer employed a high-pass filter (typically adjusted to pass only sounds above 3 kHz) to isolate high-frequency sounds and a modulating signal (typically 4.5 kHz) to transpose these sounds to lower frequencies. A low-pass filter (typically adjusted to 1.5 kHz) was included to eliminate the modulating signal. This processing shifted sound components from 3.0-4.5 kHz to 1.5-0 kHz (reversed spectrum) and from 4.5-6.0 kHz to 0-1.5 kHz. The transposed components were amplified by a compression amplifier and added to the unprocessed signal which was also amplified by a second compressor. Johansson noted that transposition resulted in new sound patterns which could prove confusing to naive listeners with normal hearing. Also, certain settings of the compressors were observed to cause consonants to mask vowels which precede them in time. In addition, some vowels were distorted by the transposition of strong high-frequency components.

Wedenberg (1959) reported on the results of training six impaired listeners in the use of the Johansson transposer for 16 hours spread over four months. Only two of the listeners appeared to have sufficient hearing to warrant speech discrimination tests before training. Training was individually paced, reflecting differences in rates of learning, and consisted of practice on consonant combinations, spondees, and connected speech. One subject, aged 29, who was severely hard of hearing from birth, improved word recognition scores from 40 to 85% using a combination of transposition and amplitude compression. No similar improvement occurred for amplification alone: scores remained constant at roughly 50%.

Johansson and Sjögren (1965) compared transposition with compression-amplification for both normals with simulated losses and hearing-impaired children. For the 12 normal listeners, high-frequency losses were simulated by a combination of 1000-Hz low-pass filtering and masking noise. Training in the use of the transposer consisted of seven 25-minute sessions spaced by weekly intervals. For small stimulus sets (4-11 CV's), large improvements in identification accuracy (from 18% to 70% correct) were observed to accompany the use of transposition, but no improvement was found when only compression amplification was provided. For large stimulus sets (phonetically balanced Swedish words), roughly equal improvements were found with transposition and amplification, although both initial and final scores favored transposition by about 10 points. The hearing-impaired children tested were 9-12 years old and had large hearing losses (80 and 95 dB at 250 and 500 Hz). Three of five children in one group were trained to achieve almost perfect identification of the syllables /sa, ∫a, and ka/ within 3-4 biweekly sessions of 5-10 minutes each using the transposer, and this training readily transferred to identification of /si, ∫i, and ki/. No improvements on this task were found

when compression amplification was used in place of the transposer. A second group of three similarly impaired children achieved comparable results on identical material.

In a later report, Johansson (1966) reported on the use of the transposer as an aid to word recognition. Six profoundly deaf (average losses of 60, 70, 85, and 100 dB at 125, 250, 500, and 1000 Hz) children were tested and trained in a series of 20-minute sessions. One of the subjects showed dramatic improvements and achieved nearly perfect recognition of a 90-word vocabulary within 3-4 sessions. The other five subjects improved from roughly 6 to 50% correct recognition with transposition within 8-10 sessions, but also improved from roughly 5 to 40% correct recognition using amplification alone (on which no training was provided).

Ling (1968, 1969) compared the Johansson transposer to his own six-channel vocoder (discussed in Section E) and to linear amplification. Eight children with residual hearing only at low frequencies (average losses of 67, 71, and 90 dB at 125, 250 and 500 Hz) were trained for 40 minutes each day over a period of 10 school days and tested in a counterbalanced order. The training, conducted by a speech-language clinician, included practice in word discrimination and articulation on an individual basis. Three tests, which used the same words on which the children had been trained, evaluated intelligibility of disyllabic words (reflecting discrimination between vowels), monosyllabic words, and monosyllabic words structured for consonant discrimination. The results, averaged over subjects, indicate that after training discrimination between words was poor (30-40% correct) and that there were no significant differences between the scores obtained with the unprocessed channel of the Johansson transposer alone, the low-frequency band of the Ling and Druz vocoder alone, the Johansson transposer with both unprocessed and transposed channels superimposed, and the six-channel Ling and Druz system. It is perhaps important to note that the sound-level output of the transposed channel of the Johansson system had been set to its maximum value and that this may have deteriorated the performance of the system. Ling concluded that his subjects were neither benefitted nor adversely affected by transposition. Ling also trained four of the children used in the above tests for 12 additional hours (over a period of 36 school days) using the Johansson transposer and materials similar to those employed in the original tests. Pre- and post-training word-identification tests were used to evaluate the Johansson device with and without the transposed channel operative. Although there was considerable variability among subjects, on the average vowel discrimination improved with training and final scores obtained with the transposed channel were slightly superior to those with amplification alone. Ling also reported that the speech-language clinician found the teaching of speech production to be facilitated by the use of vocoding or transposition even though the same children could not learn to discriminate between the sounds through hearing alone with either scheme.

Risberg (1965) constructed and tested a transposer similar to Johansson's in structure but different in some details. Risberg's device did not use compression amplifiers in either the direct or transposed channels, and inhibited transposition when low-frequency sounds were detected in the signal to be processed. The transposer was tested on 10 normal-hearing subjects who were experienced listeners but otherwise untrained on transposed speech. The processed signals were low-pass filtered sharply and masked by broadband white noise to simulate a high-frequency hearing loss. The material used in the speech tests included words and CV nonsense monosyllables. Correct-response feedback was not provided. The results were analyzed by means of confusion matrices and interpreted in terms of correct identification of place and manner of articulation. When a low-pass cutoff of 750 Hz was used, the availability of the transposed channel increased correct identification of manner of articulation from 60 to 98% but increased correct identification of place of articulation only from 33 to 44%. Nearly all the improvement in word intelligibility can be accounted for by improved identification of manner of articulation.

Ahlström, Risberg, and Lindhe (1968) evaluated the Johansson transposer on a group of 44 college students with normal hearing, making use of sharp low-pass filtering at 1000 Hz and additive broadband noise (S/N = 15 dB) to simulate high-frequency hearing loss. The subjects were divided into two groups, each of which was trained by a slightly different method. Half of each group was trained on word material and the other half on nonsense CV monosyllables. In all cases the vocabulary was limited to 48 items. Training consisted of presentation of identified transposed sounds and extended for ten 15-minute practice sessions spaced over a period of two weeks. Alternate sessions included identification tests on the practice material. The results, averaged over subjects, indicated that the performance of all four groups was about the same and improved rapidly from an initial 30% to about 60% correct identification after five sessions and 66% after all ten sessions. For all groups the primary basis for discriminating among consonants was determined to be manner of articulation, while place of articulation was not found to be identified well. Groups which practiced on word material identified manner of articulation somewhat more accurately than the groups which practiced on monosyllables. Although significant improvements were observed on tests concerned with the practice material, only very limited gains (from 20% correct to 30% correct) were made on "transfer tests" of 25 phonetically balanced meaningful words administered before and after training.

Spens and Martony (1972) have reported on an evaluation of a hearing aid incorporating transposition used in combination with lipreading. Three groups of students with normal hearing were trained, in three two-hour sessions, on a vocabulary of 50 phonetically balanced Swedish monosyllables. One group received only visual cues (lipreading) and simulated totally deaf subjects. The other two groups, who received both

visual and auditory cues, had severe impairments simulated by 54 dB/ octave lowpass filtering with a 290 Hz cutoff and background noise. The auditory signals received were either merely amplified or both transposed and amplified. The identification scores of all three groups improved by about 20 points with training. This permitted the non-auditory group nearly to double its scores and the auditory groups to achieve near-perfect scores. The magnitude of the improvement with training was also substantially larger than the differences between the scores of the two auditory groups either before or after training. An analysis of the errors made by the auditory groups after training indicated that transposition resulted in slightly (4 points) higher scores for words with voiceless stop consonants and for words with both voiceless fricatives and voiceless stops. Spens and Martony have interpreted these results as indicating that transposition may render useful a cue for distinguishing between voiceless stops and fricatives which is difficult to lipread: the duration of frication or aspiration noise.

Foust and Gengel (1973) have compared the Johansson transposer with conventional linear amplification using nine hearing impaired college students. Test materials (administered by live female voice) consisted of phonetically balanced words, and monosyllables in a Modified Rhyme Test (Kruel et al., 1968), and were presented under conditions in which lipreading could be controlled. Tests were also conducted in which only lipreading cues were available. No training was provided before testing, but correct-response feedback was given on a trial-by-trial basis, and when errors were made the experimenter and the subject interrupted the test to contrast the given and correct responses. Test sessions of one-hour duration occurred two or three times a week for a period of about five weeks in a quiet laboratory environment. During each session subjects were exposed to conventionally amplified as well as transposed sound. The results indicated wide variability over subjects and large training effects both with conventional amplification and with transposition. Averaged over subjects, transposition was found to improve performance only slightly relative to amplification: differences in scores were 10 points or less and sometimes favored conventional amplification. Confusion-matrix analyses of the errors made indicated only one systematic pattern: the phoneme /s/ was identified more correctly with transposition than with conventional amplification.

Velmans (1971, 1974) designed and patented a transposer which differs from Johansson's in that compression amplification is not used and in that frequencies from 4-8 kHz are shifted to the 0-4 kHz region. In a preliminary evaluation of this system (Velmans, 1973), 16 subjects with normal hearing with losses simulated by 900 Hz low-pass filtering and high-pass noise attempted to imitate the pronunciation of a set of 21 CVC nonsense monosyllables. Half the subjects heard signals which included transposed components; for the others the transposed channel was inoperative. All could hear their own voices after processing and filtering. In half of the

tests lipreading was permitted; in the balance it was excluded. The results indicated that transposition improves imitation performance for sounds which contain significant high-frequency components, such as fricative and stop consonants, and that this improvement occurs to about the same extent whether lipreading is present or not. Further, subjects who heard the transposed signals achieved significantly better imitation of both manner and place of articulation than those who heard only low-frequency components. However, subjects who used the Velmans transposer required about half again as much time to make a response than those who did not, indicating that the cues available in the transposed sounds could only be used with some difficulty.

Velmans (1975) has also evaluated his transposer as an aid to articulation training. Six children with pronounced sensorineural losses at high frequencies were trained to articulate monosyllabic words differing in the initial consonant (/s/, /ʃ/, /tr/, /tʃ/, /t/, and /k/) by a teacher of the deaf. Half the group was first trained (in seven 5- to 10-minute daily sessions) without transposition and then with transposition; the other half was trained in the reverse order. Evaluation, which occurred after each set of training sessions, measured correctness of articulation, and recognition of words containing the consonants but pronounced by the experimenter rather than the trainer. Although one of the children exhibited no articulation learning under either amplification condition, the group as a whole showed significant improvement. Without transposition, manner of articulation improved more than place of articulation. With transposition, both manner and place of articulation improved about the same amount. Scores on a transfer-of-training test favored transposition by 5 points for manner and 8 points for place of articulation. Recognition test results indicated substantial improvement without transposition, but essentially no change with transposition. Velmans interpreted these results as indicating that the training, although sufficient for articulation learning, was insufficient to allow the new sound patterns introduced by the transposed signals to be associated with concepts long associated with patterns conveyed solely by low-frequency speech components.

Despite the early optimistic reports by some of the proponents of transposition schemes, and despite the substantial effort which has been expended in developing and evaluating them, very little success has actually been demonstrated. The more careful and controlled studies suggest that in most cases the benefits afforded by transposer systems are very restricted. Transposition can render the high-frequency components of certain speech sounds audible for listeners with very little residual high-frequency hearing. These recoded cues may prove very useful in training speech production for selected groups of phonemes. However, the cues provided by the transposed elements generally do not enhance the recognizability or intelligibility of most speech materials. Gabrielsson et al. (1975) have recently published data on the distribution of transposed fricative energy for /s/ and /ʃ/ sounds which suggests that the distinction

between them is subtle and requires good discrimination for spectral differences in the 700-1500 Hz region. Many of the impaired listeners who participated in the evaluation of transposers may not have had adequate residual hearing to make this distinction reliably. Also, the transposed sounds may obliterate some cues normally conveyed by low-frequency speech components. This may explain in part the difference between the positive findings of studies which focused on distinctions keyed by high-frequency components (for example, Velmans, 1975) and the negative findings of studies which examined broader classes of speech sounds (for example, Ling, 1968). Although the interference of transposed and normal low-frequency cues may be unavoidable, few researchers have paid careful attention to the relative gain of the transposed channel. Finally, Foust and Gengel have pointed out that the maximum acoustic output of the commercially available body-worn transposer-aid used in some of the evaluations (124 dB SPL) may not have been adequate for the listeners used. Although this problem would likely be detected with adult subjects, it may have been ignored in some of the studies of impaired children.

H. CONCLUDING REMARKS

(1) With only minor exceptions, the results of previous research on frequency lowering have been negative. Often, the results obtained with frequency lowering have been worse than the results obtained with simple linear amplification. Furthermore, in those few cases in which frequency lowering has been superior to linear amplification, the improvement has been relatively small (and might have disappeared if certain characteristics of the linear amplification system, such as the frequency-gain characteristic, had been chosen differently).

(2) The reasons underlying the negative results obtained in previous studies of frequency lowering are essentially unknown. Among the possible explanations for the negative results are the following.

(a) The channel capacity (in the information theoretic sense) of the listeners who have been tested is too small to permit significant improvements by means of frequency lowering.

(b) There are "built-in" constraints on the speech-perception process that result in a fundamental resistance to successful recoding of speech at lower frequencies.

(c) The specific frequency-lowering displays that have been examined are inappropriate.

(d) The amount and type of training employed has been inadequate for the listeners to learn the new displays.

In our opinion, explanations (c) and (d) are more relevant than (a) and (b). Although argument (a) cannot be ruled out automatically for all impaired listeners, there exist a variety of factors which indicate that the

negative results cannot be explained in terms of reduced channel capacity per se. For example, the existence of vocoders that synthesize highly intelligible speech from low-rate control signals proves that the channel capacity of a speech communication system can be drastically reduced without substantially degrading intelligibility. Similarly, although argument (b) cannot be ruled out on the basis of existing data, we regard it as overly simplistic and of doubtful validity. Not only do listeners with normal hearing adapt readily to the spectral variations (albeit small ones) among the speech of men, women, and children, but there even exist successful methods of speech perception using nonauditory modalities. For example, there exist deaf-blind people who are known to be capable (after extensive, long-term training) of understanding conversational speech by means of tactile signals using the Tadoma Method. (Vivian, 1966; Norton et al., 1977).

To a large extent, most of the frequency-lowering schemes that have been studied were designed to achieve easy implementation on existing facilities rather than to achieve desired properties of the transformed signals. Thus, in addition to achieving some form of spectral lowering, many of these schemes altered other aspects of speech known to be important for perception. Among the characteristics that may have been unintentionally distorted are gross temporal and rhythmic patterns, fundamental frequency contours, and vowel and syllable durations. Also, many of the schemes did not preserve the normal acoustic cues for distinguishing between voiced and unvoiced sounds, a distinction that is known to carry a heavy functional load in speech communication (for example, Denes, 1963). Independent of these considerations, some of the schemes appear to have been evaluated with known distortions uncorrected or with improperly set parameters (leading to increased masking and making interpretation of results more difficult). Fortunately, the constraints imposed on frequency-lowering schemes by equipment limitations are not intrinsic. Currently available techniques, such as digital signal processing and speech synthesis, enormously expand the range of schemes that can be studied.

In addition, most of the studies that have been conducted employed only very limited training, although it is now known that the amount and type of training used may have strong effects on the perception of spectrally lowered speech. Since substantial lowering tends to create sound patterns that differ sharply from those of normal speech, it is not unreasonable to assume that such lowering can only be successful in conjunction with an intensive, long-term, and appropriately designed training program. None of the past studies pursued training significantly beyond the point at which improvement in performance first began to plateau, and most of them did not permit participants to hear their own transformed voice. In general, the amount of training that was provided was only a minute fraction of that normally provided when one attempts to learn a second language. Note also, as Erber (1971) has pointed out, that compara-

tive evaluations of frequency lowering and conventional amplification which involve children in schools for the deaf generally do not take proper account of out-of-class experience (which generally includes only conventional amplification).

Finally, most previous studies of frequency-lowering (like studies of other signal-processing schemes for the impaired) have not paid sufficient attention to the selection, categorization, and analysis of the impairments to which the frequency lowering is applied. Except for those studies that employed normal subjects with simulated losses, the residual auditory abilities of the listeners are essentially unknown (except, of course, for the absolute threshold). Furthermore, it is possible that previous research has focused too strongly on listeners with impairments that are too great. This question is particularly relevant to studies of children in schools for the deaf, where, as Erber has pointed out (Stark, 1974), there exist many cases in which the losses are so large that the cues are basically vibrotactile. Perhaps greater study of more moderate degrees of frequency lowering applied to listeners with more moderate degrees of loss would be profitable.

(3) Independent of the frequency-lowering scheme studied, the amount of training employed, and the listeners chosen as experimental subjects, it is essential that future research on frequency lowering evaluate in depth the perception of the transformed speech. On the whole, the results of previous studies provide no insight into how the transformed speech is actually perceived or into the detailed causes of the poor performance obtained. For example, there is essentially no information provided on the extent to which the poor results are attributable to inadequate resolution or to uncorrected response bias, or, given that resolution is inadequate, the extent to which poor resolution stems from basic sensory limitations or from more central limitations. Without such information it is extremely difficult to make use of the results in planning further research or even to decide whether such research is likely to prove fruitful. In our opinion the development of more effective frequency-lowering schemes depends, at least in part, on an improved understanding of why the schemes that have already been studied have not been more effective.

REFERENCES

AHLSTRÖM, K. G., RISBERG, A., and LINDHE, V., Experiment in learning to discriminate frequency transposed speech. Rpt. 36, The Institute of Education, Uppsala Univ., Sweden (1968).

BEASLEY, D. S., MOSHER, N. L., and ORCHIK, D. J., Use of frequency shifted/time compressed speech with hearing-impaired children. *Audiology*, 15, 395-406 (1976).

BENNETT, D. N., and BYERS, V. W., Increased intelligibility in the hypacusic by slow-play frequency transposition. *J. Aud. Res.*, 7, 107-118 (1967).

BENNETT, I. M., and LINVILL, J. G., A Study of time domain speech compression by means of a new analog speech processor. *J. Aud. Eng. Soc.*, 23, 713-721 (1975).

BIONDI, E., and BIONDI, L., The sampling of sounds as a new means of making speech intelligible to the profoundly deaf. *Alta. Frequenza*, 37, 180-191 (1968).

BIONDI, E., and BIONDI, L., The sampling of sounds as a means of making speech intelligible to the deaf: theories and results. Proc. 1973 Carnahan Conference on Electronic Prostheses, Univ. of Kentucky, 5-9 (1973).

BOGERT, B. P., The VOCBANC—A two to one speech bandwidth reduction system. *J. Acoust. Soc. Am.*, **28**, 399-403 (1956).

BOUILLON, J. P., Resultats de re'educations auditives effectue par des professeurs de l'Institut de Jeunes Sourds de Paris. Proc. Colloque Internat. Paris, 1967. Paris: l'Assoc. des Communautes d'Enfants, 195-196 (1968).

BREUEL, J. W., and LEVENS, L. M., The American Federation of the Blind harmonic compressor. New York City: A.F.B. (1969).

CAMPBELL, R. M., SHULTHEIS, D. C., and BARTON, J. M., A comparison of frequency spectral shifting and frequency division as a possible hearing aid. Proc. 1976 Conference on Systems and Devices for the Disabled. Tufts Univ.—New England Medical Center, Boston (1976).

DANILOFF, R. G., SHRINER, T. H., and ZEMLIN, W. R., Intelligibility of vowels altered in duration and frequency. *J. Acoust. Soc. Am.*, **44**, 700-707 (1968).

DAVID, E. E., and MCDONALD, H. S., Note on pitch-synchronous processing of speech. *J. Acoust. Soc. Am.*, **28**, 1261-1266 (1956).

DENES, P. B., On the motor theory of speech perception. In W. W. Dunn (Ed.), *Models for the Perception of Speech and Visual Form*, Cambridge: M.I.T. Press, 309-314 (1967).

DENES, P. B., On the statistics of spoken English. *J. Acoust. Soc. Am.*, **35**, 892-904 (1968).

DONE, W. J., and KIRLIN, R. L., Speech zero-crossing rate compression for bandwidth compression. *IEEE Trans. Acoustics Speech and Signal Processing*, **ASSP-23**, 433-438 (1975).

DUDLEY, H., Remaking Speech. *J. Acoust. Soc. Am.*, **11**, 169-177 (1939).

ERBER, N. P., Evaluation of special hearing aids for deaf children. *J. Speech Hearing Dis.*, **36**, 527-537 (1971).

FAIRBANKS, G., Test of phonemic differentiation: the rhyme test. *J. Acoust. Soc. Am.*, **30**, 596-600 (1958).

FAIRBANKS, G., and KODMAN, F., Word intelligibility as a function of time compression. *J. Acoust. Soc. Am.*, **29**, 636-641 (1957).

FAIRBANKS, G., EVERITT, W. L., and JAEGER, R. P., Method for time or frequency compression-expansion of speech. *IRE. Trans. Audio*, **AU-2**, 7-12 (1954).

FLANAGAN, J. L., and HOUSE, A. S., Development and testing of a formant-coding speech compression system. *J. Acoust. Soc. Am.*, **28**, 1099-1106 (1956).

FLETCHER, H., *Speech and Hearing*. New York: Van Nostrand Co. (1929).

FLETCHER, H., *Speech and Hearing in Communication*. Princeton: Van Nostrand (1953).

FOULKE, E., and STICHT, T. G., Review of research on the intelligibility and comprehension of accelerated speech. *Psych. Bull.*, **72**, 50-62 (1969).

FOUST, K. O., and GENGEL, R. W., Speech discrimination by sensorineural hearing impaired persons using a transposer hearing aid. *Scand. Audiol.*, **2**, 161-170 (1973).

FRENCH, N. R., and STEINBERG, J. C., Factors governing the intelligibility of speech sounds. *J. Acoust. Soc. Am.*, **19**, 90-119 (1947).

GABRIELSSON, A., JOHANSSON, B., LINDBLAD, A. C., PERSSON, L., PETTERSSON, A., and ROSENQVIST, B., Frequency discrimination for bands of noise. *Audiology*, **14**, 1-20 (1975).

GUTTMAN, N., LEVITT, H., and BELLEFLEUR, P. A., Articulation training of the deaf using low-frequency surrogate fricatives. *J. Speech Hearing Res.*, **13**, 19-29 (1970).

GUTTMAN, N., and NELSON, J. R., An instrument that creates some artificial speech spectra for the severely hard of hearing. *Am. Ann. Deaf*, **113**, 295-302 (1968).

GUTTMAN, N., and VAN BERGEIJK, W. A., Frequency division in real time as a means of counteracting high frequency hearing loss. Societe Internat. D'Audiol., Fourth Congress, 1958, Lyon: Audin (1959).

HASPIEL, G. S., Speech compression and frequency translation for hearing-impaired auditors. Tech. Prog. Rpt., Proj. HEW SRS, RD 2575-S (1969).

HOPNER, E., and ANDREWS, M. C., Consonant response in narrow band transmission. U.S. Patent 3,069,506 (1962).

JOHANSSON, B., A new coding amplifier system for the severely hard of hearing. Proc. 3rd International Congress on Acoustics, Stutgart, 655-657 (1961).

JOHANSSON, B., The use of the transposer for the management of the deaf child. *Internat. Audiol.*, **5**, 362-373 (1966).

JOHANSSON, B., and SJOGREN, H., The effect on normal hearing subjects and profoundly deaf children by coding voiceless consonants using frequency transposition. Proc. 5th International Congress on Acoustics, Liege, 1-4 (1965).

KLUMPP, R. G., and WEBSTER, J. C., Intelligibility of time-compressed speech. *J. Acoust. Soc. Am.*, **33**, 265-267 (1961).

KNORR, S. G., A hearing aid for subjects with extreme high frequency losses. *IEEE Trans. Acoust., Speech, Signal Processing*, **ASSP-24**, 473-480 (1976).

KRUEL, E. J., NIXON, J. C., KRYTER, K. D., BELL, S. W., LANG, J. S., and SCHUBERT, E. D., A proposed clinical test of speech discrimination. *J. Speech Hearing Res.*, **11**, 536-552 (1968).

KURTZROCK, G., The effects of time and frequency distortion upon word intelligibility. Unpublished Ph.D. thesis, Univ. of Illinois (1956).

LAFON, J. C., Appareil amplificateur a compensation pour la surdite profonde. *J. Internat. Audiol.*, **6**, 343-345 (1967).

LEE, F. F., Time compression and expansion of speech by the sampling method. *J. Aud. Eng. Soc.*, **20**, 738-742 (1972).

LING, D., Three experiments on frequency transposition. *Am. Ann. Deaf*, **113**, 283-294 (1968).

LING, D., Speech discrimination by profoundly deaf children using linear and coding amplifiers. *IEEE Trans. Aud. Electroacoust.*, **AU-17**, 298-303 (1969).

LING, D., Auditory discrimination of speech transposed by a sample-and-hold process. In G. Fant (Ed.), *Proc. International Symposium on Speech Communication and Profound Deafness, Stockholm (1970)*. Washington: A. G. Bell Assoc. (1972).

LING, D. and DOEHRING, D. G., Learning limits of deaf children for coded speech. *J. Speech Hearing Res.*, **12**, 83-94 (1969).

LING, D. and DRUZ, W. S., Transposition of high-frequency sounds by partial vocoding of the speech spectrum: its use by deaf children. *J. Aud. Res.*, **7**, 133-144 (1967).

LING, D. and MARETIC, H., Frequency transposition in the teaching of speech to deaf children. *J. Speech Hearing Res.*, **14**, 37-46 (1971).

MAZOR, M., SIMON, H., SCHEINBERG, J., and LEVITT, H., Moderate frequency compression for the moderately hearing impaired. *J. Acoust. Soc. Am.*, **62**, 1273-1278 (1977).

MILLER, G. A., and LICKLIDER, J. C. R., The intelligibility of interrupted speech. *J. Acoust. Soc. Am.*, **22**, 167-173 (1950).

MINAMI, K., and KATSURAYAMA, K., Frequency shift type hearing aid and an utterance trainer. *Am. Ann. Deaf*, **113**, 307-310 (1968).

MORGAN, H. L., and SAVIDGE, J. J., Bandwidth compression and expansion. M. S. Laboratories, P. O. Box 28425, Dallas, Texas (1975).

NAGAFUCHI, M., Intelligibility of distorted speech sounds shifted in frequency and time in normal children. *Audiology*, **15**, 326-337 (1976).

NIELSEN, T. E., Information about transposition. Paper presented at the 2nd Oticongress, Copenhagen (1972).

NORTON, S., SCHULTZ, M. C., REED, C. M., BRAIDA, L. D., DURLACH, N. I., RABINOWITZ, W. M., and CHOMSKY, C., Analytic study of the Tadoma method: background and preliminary results. *J. Speech Hearing Res.*, **20**, 574-595 (1977).

OCHIAI, Y., and IZUMITACHI, N., Timbre study on mishearing phenomena of speech phones in rotational synchronous distortion. Report I: Mishearing vowelic timbre. *Mem. Fac. Eng. Nagoya*, **7**, 49-60 (1955).

OCHIAI, Y., SAITO, S., and SAKAI, Y., Articulation Study of speech qualities in rotational synchronous distortion. *Mem. Fac. Eng. Nagoya*, **7**, 40-48 (1955).

OEKEN, F. W., Can hard of hearing patients suffering from high-tone perceptive deafness be improved by frequency transposition? *J. Internat. Audiol.*, **2**, 263-266 (1963).

PICKETT, J. M., Status of speech-analyzing communication aids for the deaf. *IEEE Trans. Aud. Electroacoust.*, **AU-20**, 3-8 (1972).

PIMONOW, L., The application of synthetic speech to aural rehabilitation. *J. Aud. Res.*, **3**, 73-82 (1963).

PIMONOW, L., Technical and physiological problems in the application of synthetic speech to aural rehabilitation. *Am. Ann. Deaf*, **113**, 275-294 (1968).

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

RAYMOND, T. H., and PROUD, G. O., Audiofrequency conversion. *Arch. Otolaryng.*, **76**, 60-70 (1962).

REDDEN, R. B., The effects of frequency compression and transposition on the intelligibility of monosyllabic words in normal hearing listeners. Unpublished E.D. thesis, Boston Univ. (1973).

REEDER, K., A low-frequency formant speech code for the hearing impaired. Unpublished Ph.D. thesis, Brigham Young Univ. (1975).

REEDER, K., STRONG, W. J., and PALMER, E. P., Low-frequency formant speech code for the severely/profoundly hearing impaired. *J. Acoust. Soc. Am.*, **58**, S113(A) (1975).

RISBERG, A., The transposer and a model for speech perception. STL-QPSR, 4, 26-30, KTH, Stockholm (1965).

RISBERG, A., A critical review of work on speech analyzing hearing aids. *IEEE Trans. Aud. Electroacoust.*, **AU-17**, 290-297 (1969).

SCHREINER, P. G., A speech processing aid for persons with profound sensorineural hearing loss: an application of bandwidth compression and dynamic range reduction. Unpublished Ph.D. thesis, Univ. of Michigan (1977).

SCHROEDER, M. R., Vocoders—analysis and synthesis of speech. *Proc. IEEE*, **54**, 720-734 (1966).

SCHROEDER, M. R., FLANAGAN, J. L., and LUNDRY, E. A., Bandwidth compression of speech by analytic-signal rooting. *Proc. IEEE*, **55**, 396-401 (1967).

SCHROEDER, M. R., LOGAN, B. F., and PRESTIGIACOMO, A. J., New methods for speech analysis-synthesis and bandwidth compression. Paper G8, Speech Communication Seminar, Stockholm, 1-14 (1972).

SCOTT, R. J., and GERBER, S. E., Pitch-synchronous time compression of speech. Proc. 1972 Conference on Speech Communication and Processing, IEEE Cat. No. 72 CHO 596-7 AE, 63-65 (1972).

SEARLE, J. L., Development of a portable speech compressor for the hard of hearing. Final Report: HRB Singer, Inc. (1969).

SHRINER, T. H., BEASLEY, D. S., and ZEMLIN, W. R., The effects of frequency division on speech identification in children. *J. Speech. Hearing Res.*, **12**, 413-422 (1969).

SPENS, K. E., and MARTONY, J., The transposer as a lipreading aid. Proc. 1972 Conference on Speech Communication and Processing, IEEE Cat. No. 72 CHO 596-7 AE, 246-248 (1972).

STARK, R., *Sensory capabilities of hearing-impaired children.* Baltimore: Univ. Park Press (1974).

STEWART, S. E., STRONG, W. J., and PALMER, E. P., Experiments on the intelligibility of speech codes for the severely hearing impaired. *J. Acoust. Soc. Am.*, **60**, Suppl. 1. S124 (1976).

STOVER, W. R., Electronic speech processing methods for the hearing impaired. *J. Aud. Res.*, **7**, 313-325 (1967).

STRONG, W. J., and PALMER, E. P., Formant frequency estimation from linear prediction spectra and their second derivatives. *J. Acoust. Soc. Am.*, **55**, 396(A) (1974).

TAKEFUTA, Y., and SWIGART, E., Intelligibility of speech signals spectrally compressed by a sampling-synthesizer technique. *IEEE Trans. Aud. Electroacoust.*, **AU-16**, 271-274 (1968).

THOMAS, I. B., and FLAVIN, F. E., The intelligibility of speech transposed downward in frequency by one octave. *J. Aud. Eng. Soc.*, **18**, 56-62 (1970).

TIFFANY, W. R., and BENNETT, D. N., Intelligibility of slow played speech. *J. Speech Hearing Res.*, **4**, 248-258 (1961).

TOONG, H. D., A study of time-compressed speech. Unpublished Ph.D thesis, Massachusetts Institute of Technology (1974).

VELMANS, M., Aids for deaf persons. U. K. Patent 1340105 (1971).

VELMANS, M., Speech imitation in simulated deafness, using visual cues and recoded auditory information. *Language and Speech*, **16**, 224-236 (1973).

VELMANS, M., The design of speech recoding devices for the deaf. *Brit. J. Audiol.*, **8**, 1-5 (1974).

VELMANS, M., Effects of frequency recoding on the articulation learning of perceptively deaf children. *Language and Speech*, **18**, 180-193 (1975).

VIVIAN, R. M., The Tadoma method: a tactual approach to speech and speech reading. *Volta Rev.* (Dec.), 733-737 (1966).

VOIERS, W. D., SHARPLEY, A. D., and HELMSOTH, C. J., Research on Diagnostic evaluation of speech intelligibility. Austin: TRACOR Inc. (1973).

WEDENBERG, E., Auditory training of the severely hard of hearing using a coding amplifier. Proc. 3rd International Congress on Acoustics, Stuttgart, 658-660 (1959).

ZEMLIN W. R., The use of bandwidth and time compression for the hearing handicapped. In E. Foulke (Ed.), Proc. of the Louisville Conference on Time Compressed Speech. Louisville: Univ. Louisville Press, 95-97 (1966).

Postface

Generally speaking, the research reviewed in Chapters II-IV is far from inspiring. Not only has this research failed to provide substantial improvements in speech perception for impaired listeners, but it has not greatly increased understanding of the fundamental limitations on achieving this goal. Theoretically at least, it should be possible either to develop schemes that radically improve speech perception or to obtain a fundamental understanding of why such schemes cannot be developed.

Clearly, a major obstacle to advances in this field is the intrinsic complexity of the problem and lack of fundamental knowledge in the areas of speech perception and sensorineural hearing impairments. Speech perception, even in normals, is not well understood. Also, relatively little is known (at both the physiological and perceptual levels) about the auditory systems of the specific impaired listeners who serve as experimental subjects in the tests of the various schemes. This is due in part to inadequate background knowledge on impaired hearing and in part to the problems associated with characterizing the abnormalities of a given listener. Symptomatic of this state of affairs is the paucity of theoretical models of impaired hearing (or of serious attempts to determine theoretical upper bounds on the improvements that are obtainable with signal processing of any kind). The often-stated question concerning why hearing aids are so inferior to eyeglasses is inappropriate: the problems solved by eyeglasses are analogous to conductive problems, not sensorineural problems.

Advances have also been impeded by the failure to mount adequate research programs on the matching of speech to residual auditory function. One such inadequacy concerns the character of the research staffs involved: they generally have not been sufficiently interdisciplinary. An ideal research program would include speech and hearing scientists, clinicians, engineers, psychoacousticians, and experts in perception and perceptual learning. Many of the deficiencies in past studies could easily have been avoided had the knowledge and sensitivities implied by such a staff been available to the research. A second important inadequacy concerns the style of the research: it generally has not been sufficiently long-term or sufficiently analytic. An ideal program would rigorously compare a wide variety of potentially useful signal processing schemes; vary the relevant parameters of the different signal-processing schemes systematically; employ a comprehensive set of speech materials and listening conditions; specify the signal-processing schemes, the speech materials, and listening conditions accurately and in detail; and study individual im-

114

paired listeners in depth. Also, and perhaps most important, it would attempt to determine the underlying causes of the results obtained, independent of whether these results are positive or negative. Without such determination, it is extremely difficult to build on the results of past research.

Finally, we wish to note that the topics considered in this review by no means exhaust the set of all topics in the area of matching speech to residual auditory function. Generally speaking, the schemes that have been considered focus on the problem of reduced auditory area; they do not attempt to combat distortions within the auditory area. (The principal exception to this limitation concerns the use of amplitude compression to combat recruitment.) Furthermore, the review has focused on monaural hearing; it has essentially ignored the fact that most impaired listeners have some residual hearing in both ears. Clearly, the optimal signal-processing scheme for such listeners must take account of the hearing in both ears and of binaural interaction. There is no reason to believe that the optimum binaural aid is composed of two monaural aids that are optimum for the two ears considered separately, or that the monaural aid that is optimum when the second ear has no input will be optimum when the second ear is stimulated.

Additional important topics that we have not considered (aside, of course, from those involving other sense modalities), concern the use of schemes that employ special instrumentation on the talker (for example, schemes in which the talker wears a microphone to enhance the signal-to-noise ratio) and schemes that employ electrical stimulation (for example, cochlear implants). Although the first class of schemes may be of great value in special situations, they do not provide a general substitute for schemes that do not impinge on the talker. Similarly, although the second class of schemes can provide a "sense of hearing" to some patients who were formerly deaf, it is unlikely that speech perception can be radically improved by these schemes without substantial signal processing at the input. Independent of whether a listener's limited hearing capacity results from a partial natural impairment or from a very severe natural impairment that has been partially alleviated by the use of electrical stimulation, one is still faced with the basic problem of developing signal-processing schemes to match speech to residual auditory functions. To date, however, very little attention has been given to this problem in the context of electrical stimulation.