This is to certify that the

thesis entitled

QUALITY MAGNITUDE ESTIMATION FUNCTIONS FOR
DEGRADED SPEECH BY LISTENERS WITH NORMAL HEARING AND
LISTENERS WITH SENSORINEURAL HEARING LOSS

presented by

Gary Dean Lawson

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Audiology and
Speech Sciences

Michael R. Chial, Ph.D.
Major professor

Date February 7, 1980

~~207~~

201

QUALITY MAGNITUDE ESTIMATION FUNCTIONS FOR DEGRADED SPEECH BY LISTENERS WITH NORMAL HEARING AND LISTENERS WITH SENSORINEURAL HEARING LOSS

By

Gary Dean Lawson

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Audiology and Speech Sciences

1980

ABSTRACT

QUALITY MAGNITUDE ESTIMATION FUNCTIONS FOR
DEGRADED SPEECH BY LISTENERS WITH NORMAL HEARING
AND LISTENERS WITH SENSORINEURAL HEARING LOSS

By

Gary Dean Lawson

Although traditional word discrimination tests continue to be widely used in clinical settings, they do not adequately predict listener performance in everyday life. Speech quality judgments may be helpful in this respect, since they have sometimes differentiated among hearing aids when word discrimination tests did not. Insufficient research is available, however, to justify the clinical use of speech quality judgments on a routine basis. Clinical research on quality judgments has, for the most part, employed the method of paired comparisons. Other methods, for example direct magnitude estimation, have received little or no attention.

This study investigated speech quality magnitude estimates (SQMEs) by 12 normal hearing listeners (Group 1) and 12 sensorineurally impaired hearing listeners (Group 2) as a function of seven degrees of three degradation types (low-pass filter bandwidth, high-pass filter bandwidth, and percent total harmonic distortion by linear rectification.) The purpose was to determine: (1) whether the psychophysical power law applies to the scaling of speech quality and (2) whether there are differences in the log SQME - log degree of degradation functions as a function of listener group, degradation type, and listener group-by-degradation type interaction.

Prior to the listening tasks, each subject participated in visual magnitude estimation training and screening tasks. Dependent variables were (1) visual magnitude estimates of circle size and (2) slopes of the least squares lines of best fit which related log visual magnitude estimates to log circle size. The visual magnitude estimates and the slopes showed excellent within-session repeatability. The log-log functions were relatively linear and showed roughly equivalent mean slopes (about 0.7) for Groups 1 and 2. Both groups reliably produced expected data and appeared to have similar visual magnitude estimation skills.

Dependent variables for the listening tasks included (1) log geometric mean SQMEs across trials for each degree of each degradation type and (2) the slopes of the least squares lines of best fit for the log SQME - log degree degradation functions. A subgroup of four subjects in each group repeated the listening tasks in a second session. Between-session reliability of log geometric mean SQMEs for individual subjects was very high for both subgroups under each degradation type, but between-session reliability of slopes for the two subgroups showed considerable variability as a function of group-by-degradation type interaction. Log geometric mean SQMEs for Groups 1 and 2 increased linearly as a function of decreasing log degree of degradation. The slopes of the log-log functions differed as a function of degradation types and group-by-degradation type interaction.

The excellent reliability of the visual magnitude estimation data suggests that systematic differences in performance on the SQME tasks are probably due to perceptual differences. The linear relationship between log geometric mean SQMEs and log degree of each degradation type indicates that a power function exists in each case. Systematic slope differences among the log-log functions were attributed to perceptual differences. Estimates of poor between-session reliability of slopes were attributed to perceptual difficulties. Collectively, the findings were sufficiently encouraging to warrant additional research. Possible areas of research include the application of SQMEs to the evaluation of communication systems, clinical practices in audiology, and how normal hearing and hearing impaired individuals process complex signals.

To Becky

ACKNOWLEDGEMENTS

Appreciation is extended to the members of my dissertation guidance committee: Michael R. Chial, Herbert J. Oyer, Joseph R. Vorro, and Steven C. White. Special gratitude is owed to Michael R. Chial, who served as the dissertation advisor, provided extensive help in the development of the visual training task, provided perceptive comments on the manuscript, and gave support and encouragement during hard times. Special thanks are also due to Herbert J. Oyer for guidance through most of my academic program and for encouragement of early attempts at research.

A debt is owed to the subjects who volunteered for the study and to a number of fellow students, friends, and associates. Michael L. Stouffer spent many hours, days, and nights writing computer programs; Lynn Waters was the talker for the experimental stimuli; Linda L. Smith provided materials for data collection; and Carol Goldschmidt and Nancy Brewer assisted in selecting stimulus materials. Thanks go to my colleagues at Western Michigan University for their patience and understanding and to the Western Michigan University Computer Center for assistance in analyzing the data.

The person who deserves the most thanks is my wife, Becky, for her interminable patience, understanding, encouragement, and loving help through it all. Special thanks go to our parents for their continued support and encouragement and for understanding when we weren't there.

TABLE OF CONTENTS

CHAPTER PAGE

CHAPTER		PAGE

CHAPTER PAGE

CHAPTER		PAGE

LIST OF TABLES

TABLE		PAGE

TABLE		PAGE

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Several reviewers (e.g., Chial and Hayes, 1974; Oyer and Frankman, 1975; Millin, 1975; and Berger, 1978) have indicated that traditional word discrimination tests do not predict a listener's communicative effectiveness in the real world and therefore do not measure handicap. Communicative effectiveness in the real world may be more closely related to the magnitude of "goodness" or overall quality of the speech a listener perceives than to the mere intelligibility of it. Licklider (1946) concluded that "amplitude distortion affects quality somewhat more severely than it does intelligibility" (p. 432). This is not to say, however, that quality and intelligibility are unrelated. Weldele (1973) and Weldele and Millin (1975) reported a significant relationship between preference-based ratings and discrimination-based ratings of hearing aids. It seems reasonable to assume that intelligibility contributes to the "goodness" or overall quality of one's perception of speech.

Although there appears to be some interest in the use of speech quality judgments as a clinical tool, clinical methods based upon quality judgments have not been well researched. In spite of the frequent complaints about traditional monosyllabic word tests, they continue to be widely used in clinical settings (Burney, 1972; Martin and Pennington, 1971; Martin and Forbis, 1978). Although Weldele and Millin (1975) encouraged the use of quality judgments in hearing aid evaluations, they did not suggest that the use of discrimination tests be discontinued. As noted by Punch (1978), the difficulties with traditional methods dictate that basic procedural issues and assumptions be carefully evaluated before audiologists adopt any new clinical strategies. The first step is to examine what has already been done.

Background

Although there are a number of approaches to the assessment of speech quality, clinical research has emphasized the method of paired comparisons. Relatively little attention has been given to other psychophysical methods (e.g., magnitude estimation) or to theoretical considerations.

General Approaches to Speech Quality Measurement

The work on speech quality measurement was surveyed by Munson and Karlin (1962) and by Hecker and Guttman (1967). Munson and Karlin divided methods of measurement into "indirect comparisons" by which transmission systems are assessed singly and "direct comparisons" by which systems are assessed in pairs, as in paired comparisons. Hecker and Guttman (1967) categorized methods as (1) analytic approaches which aim to discover the psychological attributes of the speech signal and (2) utilitarian approaches which are concerned with determining speech quality by prior assumption of psychological attributes and reduction of measures to a unidimensional scale.

The Institute of Electronics and Electrical Engineers (IEEE, 1969) recommended procedures for using subjective preference measurements to estimate speech quality. Utilitarian methods were said to be best suited for engineering practice, and three such methods were outlined: (1) the Isopreference Method, (2) the Relative Preference Method, and (3) the Category-Judgment Method. The *Isopreference Method* involves the comparison of a test signal to a referent signal subjected to varying degrees of degradation. The isopreference level is the signal-to-noise ratio of the test and reference signal at which the test and reference signals are preferred an equal number of times. The Relative Preference Method seeks to determine the quality of the test signal by locating it on a quality continuum, which is defined by reference signals representing different types of speech distortion. The test signal is positioned on the continuum by considering how often it is preferred to any reference signal. In the Category-Judgment Method listeners describe their impression of the quality of a speech signal by assigning it to one of several simple categories (e.g., Unsatisfactory, Poor, Fair, Good, Excellent).

The study of speech quality by audiologists has involved a "utilitarian" approach to the evaluation of signals transduced by hearing aids. Apparently, every study to date has employed a direct paired-comparison paradigm.

Paired-Comparison Quality Judgments

The first experiment on the clinical use of speech quality judgments was a paired-comparison study by Jeffers (1960). Today, at least eight studies have included paired-comparison quality judgments.

Three studies included only hearing-impaired listeners. Jeffers (1960) asked 32 subjects with conductive hearing losses to give preferences for the quality of speech transduced by five hearing aids arranged in pairs. Zerlin (1962) asked 21 subjects with sensorineural hearing losses to state preferences for the quality of speech transduced by six hearing aids arranged in pairs. Weldele and Millin (1975) obtained preference judgments from 10 listeners with sensorineural hearing loss on pairings of four hearing aids.

Three additional studies included only normal hearing listeners. Witter and Goldstein (1971) obtained quality preference judgments from 30 normal hearing listeners on pairings of five hearing aids. Smaldino (1974) obtained quality preferences from 10 normal hearing subjects on pairings of stimuli transduced by 10 hearing aids. Yonovitz, Bickford, Lozar, and Ferrell (1978) obtained paired-comparison judgments and dissimilarity ratings of 12 hearing aids from 20 normal hearing listeners.

Each of the two remaining studies included a group of listeners with normal hearing and a group of listeners with sensorineural hearing loss. Punch and Ciechanowski (1977) obtained preferences from 10 subjects in each group on pairings of stimuli transduced by five hearing aids. Chial and Daniel (1977) obtained preferences from 18 subjects in each group, using a magnitude estimation procedure as well as a paired comparison procedure on stimuli transduced by four hearing aids.

It is helpful to examine the paired-comparison studies in terms of the sensitivity of quality judgments to electroacoustic and other stimulus characteristics, quality judgment reliability, and feasibility issues.

Sensitivity to electroacoustic characteristics. Several studies have suggested that paired-comparison quality judgments are related to the electroacoustic characteristics of hearing aids (e.g., Jeffers, 1960; Zerlin, 1962; Witter and Goldstein, 1971; Smaldino, 1974; and Yonovitz *et al*., 1978). In general, these studies found that hearing aids with better electroacoustic characteristics are preferred over those with poorer electroacoustic characteristics.

Jeffers (1960) and Zerlin (1962) found that preference tests differentiated among hearing aids when monosyllabic word discrimination tests did not. Unfortunately, neither investigator measured the electroacoustic characteristics of the experimental hearing aids. Although Jeffers categorized aids on the basis of the manufacturer's specifications, Kasten and Revoile (1965) showed that the actual electroacoustic characteristics of hearing aids may differ significantly from the manufacturer's design specifications. Apparently, Zerlin simply assumed that the differences in preferences for hearing aids were due to differences in electroacoustic characteristics.

Weldele and Millin (1975) found that when hearing aids were rated by word discrimination scores obtained at 40 dB HTL, discrimination-based ratings and preference based ratings were significantly related. The authors reported only general descriptions of the frequency response and gain characteristics of the hearing aids. They did not comment specifically on the relationship of quality judgments and electroacoustic characteristics.

Witter and Goldstein (1971) and Smaldino (1974) were more systematic in that they measured the electroacoustic characteristics of their hearing aids. Considering male and female voice stimuli separately, Witter and Goldstein found strong Spearman rank-order correlations (0.60 to 1.00) among modal preference rankings for hearing aids and rankings of transient response, frequency range, and harmonic distortion measures. Preference judgments were correlated positively with the high cut-off point for frequency response and negatively with the low cut-off point. Very low correlations were observed for preference judgments and intermodulation distortion. Smaldino (1974) found that quality judgments were (1) negatively correlated (Pearson product-moment correlations) with harmonic distortion measures at 800 Hz (-0.56), 1200 Hz (-0.59), and 1600 Hz (-0.54) with an input at 400 Hz and (2) positively correlated with the Houston Speech and Hearing Center (H.S.H.C.) bandwidth above 1000 Hz

(0.53). The H.S.H.C. bandwidth (Jerger and Thelin, 1968) is determined by drawing a line parallel to the frequency axis 10 dB below the highest point on a hearing aid response curve. The upper limit is the frequency at which the line intersects the curve above 1000 Hz; the lower limit is the frequency at which the line intersects the curve below 1000 Hz. Smaldino examined bandwidth above 1000 Hz, below 1000 Hz, and total bandwidth. Three other studies, Punch and Ciechanowski (1977), Chial and Daniel (1977) and Yonovitz et al. (1978), also included measures of electroacoustic characteristics. Although Punch and Ciechanowski reported listener preferences among the five aids employed, they did not comment specifically on the relationship of quality judgments to electroacoustic characteristics. In the Chial and Daniel study, correlations among various electroacoustic measures and quality judgments were not significant; this finding was interpreted as failing to confirm or refute previous claims about the sensitivity of quality judgments to measurable electroacoustic differences. Yonovitz et al. (1978) found that: (1) frequency response and bandwidth affected the perception of speech and music, (2) third harmonic distortion and internal noise were specific to speech perception, and (3) transient distortion and phase distortion were specific to music perception.

Some insight into the sensitivity factor in subjects with sensorineural hearing loss may be gained from examining reliability data. For example, Punch and Ciechanowski (1977) found that the quality preferences of normal and sensorineurally impaired listeners were highly correlated (Pearson r = 0.98), but noted that fewer hearing-impaired listeners were able to replicate their first-preference judgments. Chial and Daniel (1977) reported that although normal and dysacusic listeners expressed similar preferences for better quality signals, dysacusic listeners were less consistent in expressing preferences when overall signal quality was low. These results suggest that the sensitivity of quality judgments to differences in electroacoustic characteristics may vary as a function of hearing acuity. This cannot be confirmed, however, on the basis of previous research. Jeffers (1960) failed to measure electroacoustic characteristics and used only conductively impaired listeners who probably responded much like normal listeners. Zerlin (1962) used sensorineurally impaired listeners, but failed to measure the electroacoustic characteristics of his hearing aids. The more systematic studies of Witter and Goldstein (1971) and Smaldino (1974) included only normal listeners.

Sensitivity to other stimulus characteristics. Studies of speech quality have taken a number of approaches to the selection of stimulus materials. Jeffers (1960, p. 261) used eight one-minute tape recorded paragraphs "...believed to have little innate appeal..." The source of the paragraphs and the sex of the speaker were unidentified. Zerlin (1962) used 30-second passages of *Reader's Digest* material tape recorded by a speaker whose dialect was General American and whose sex was unidentified. Witter and Goldstein (1971) used a single 10-second paragraph from a Thurber short story. The passage was tape recorded by a male and a female speaker of unspecified dialect. Smaldino (1974) also used a single passage. Weldele and Millin (1975) used word discrimination lists (CID Auditory Test W-22). In addition to music stimuli, Punch and Ciechanowski (1977) used 30-second passages from Mark Twain's *Tom Sawyer* recorded by a male and a female speaker, while Yonovitz *et al*. (1978) used a 30-second paragraph from *The Rainbow Passage* which was read by a male speaker of the General American dialect. Chial and Daniel (1977) used four passages tape recorded by a female speaker of General American dialect. These passages, originally intended for reading by high school and junior high students, were previously used in Chial's (1973) dissertation. Chial judged these passages to be "...approximately equivalent along the dimensions of abstraction, grammatical complexity, vocabulary, intrinsic interest, and controversiality" (p. 19).

Witter and Goldstein (1971), Smaldino (1974), and Yonovitz et al. (1978) attempted to control for differences in stimulus materials by using a single passage for all measurements. Chial (1973) judged his stimuli to be approximately equivalent along certain linguistic and literary dimensions. On the other hand, Jeffers (1960) used no discernible control. Apparently, Zerlin (1962) and Punch and Ciechanowski (1977) assumed their passages to be equivalent in reading difficulty. The passages were presented, however, as listening rather than reading tasks, and the degree to which readability predicts listenability is unclear (Klare, 1963). Although no study has investigated the sensitivity of quality judgments to differences in stimulus materials, some have failed to establish approximate equivalency of stimulus materials.

The effects of talker sex differences upon quality judgments were discussed by Witter and Goldstein (1971) and Punch and Ciechanowski (1977). Witter and Goldstein reported differences in the proportion of preferences for a given hearing aid as a function of whether the stimuli were presented by a female voice or a male voice. It was noted, however, that the proportion of preferences also depended upon the particular pair of hearing aids being compared, raising the issue of hearing aid-voice interaction. Punch and Ciechanowski (1977) reported statistically significant ($p < 0.05$) Pearson product-moment correlations between quality preferences for (1) male and female voices (0.72 for normal listeners and 0.89 for sensorineurally impaired listeners), (2) male voice and music (0.92 for normal listeners and 0.89 for sensorineurally impaired listeners), and (3) female voice and music (0.89 for normal listeners and 0.94 for sensorineurally impaired listeners). Each correlation was said to account for an acceptable proportion of variance, and it was concluded that overall hearing aid-stimulus interaction was absent.

Reliability. Jeffers (1960) concluded that quality judgments made by conductively impaired listeners are reliable. However, this conclusion was based upon intersubject agreement rather than test-retest agreement. In this case intersubject agreement probably represents the sensitivity of quality judgments to differences in hearing aid performance more than the repeatability of quality judgments. Zerlin (1962) performed no statistical test of paired-comparison judgment reliability. However, on the basis of a table of test-retest comparisons of hearing aid preference ranks by individual sensorineurally impaired listeners, he noted that reliability appeared to be "encouraging." Witter and Goldstein (1971) and Yonovitz et al. (1978) did not report specific reliability data. Witter and Goldstein (1971) did note that a partial replication of their experiment yielded similar results. Weldele and Millin (1975) also did not report test-retest data.

Smaldino (1974) reported a fairly strong Pearson product-moment correlation ($r = 0.827$) between test-retest quality judgments, indicating good reliability with normal hearing listeners. Punch and Ciechanowski (1977) obtained statistically significant ($p < 0.05$) Pearson product-moment correlation coefficients for the test-retest comparisons made by the normal hearing listeners and the dysacusic listeners on each stimulus condition. The reliability coefficients for the normal hearing group were: 0.86 for the male voice, 0.69 for the female voice, and 0.56 for the music stimuli; coefficients for the dysacusic listening group were: 0.85 for the male voice, 0.54 for the female voice, and 0.35 for the music.

The authors concluded that only the male voice stimulus resulted in acceptably reliable data for clinical purposes with dysacusic patients. Chial and Daniel (1977) reported test-retest correlation coefficients (Spearman rank order correlations) of 0.98 for normal listeners and 0.94 for dysacusic listeners.

Feasibility issues. The use of paired-comparison quality judgments to determine the rank order of hearing aids can be an excessively time consuming affair. Punch and Ciechanowski (1977) indicated that reliability may be affected by the talker, while Chial and Daniel (1977) suggested that the judgments of dysacusic listeners are less consistent when signal quality is low. Thus, it is probably wise to obtain multiple observations for each paired comparison. In addition, the determination of rank order required the comparison of each aid with every other aid.

Quality Magnitude Estimation

One alternative to paired comparisons is quality magnitude estimation. Magnitude estimation is a procedure by which the quality of signals may be assessed singly. The listener's task is to assign a numerical value to each of several signals. Sometimes the stimulus set includes a preselected reference signal to which a specific value (i.e., a modulus) has been assigned. In other cases the listener might simply be asked to assign his own numerical value to a given stimulus and then rate the other stimuli relative to that one. Magnitude estimation is a procedure suggested by S. S. Stevens (1957) for obtaining ratio data. According to Stevens, discrimination of sensations can be accomplished by less stringent scaling methods (e.g., ordinal scaling), but sensation magnitude is directly and validly measured only by ratio scaling.

Chial and Daniel (1977) employed a quality magnitude estimation procedure which included three preselected reference signals. Since the listeners used a graphic equal-appearing interval scale to assign integer values from zero to 10 to each of the stimuli, they produced interval data. Chial and Daniel compared the reliability and sensitivity of a paired-comparison method and their "magnitude estimation method" for measuring the quality of hearing aid transduced speech as perceived by normal and dysacusic listeners. It was found that both methods produced reliable data. Normal and dysacusic listeners expressed similar preferences and similar quality magnitude estimates for better quality signals, but when signal quality was low, dysacusic listeners were obviously less consistent in expressing preferences and in estimating quality magnitude. Correlations among various electroacoustic measures and group performance on the two quality measurement tasks were not significant; this finding was interpreted as failing to confirm or refute claims about the sensitivity of the paired comparison and magnitude estimation methods to measurable electroacoustic differences. However, it was suggested that direct quality magnitude estimation procedures may provide information useful to the understanding of how impaired listeners process complex signals.

Quality judgments represent psychological responses to physical stimuli. The problem is one of discovering a simple equation which describes the relationship between the physical parameters of stimuli and subjective reactions to them. S. S. Stevens (1975) reviewed many of his earlier magnitude estimation experiments in which psychophysical functions were often displayed as straight lines when both the sensation and stimulus magnitudes were plotted as logarithmic coordinates. Stevens' data conformed to a straight line equation,

$$\log \psi = \beta \log \phi + \log \kappa$$

where ψ represents the estimated psychological magnitude and ϕ represents the stimulus magnitude. Beta (β) represents the slope of a line and κ is a constant. Changes in different physical stimuli lead to different subjective reactions and therefore, to different beta terms. Thus, beta becomes a dependent variable suggesting differences in perceptual events for different physical parameters. Taking the antilogarithms, the equation becomes

$$\psi = \kappa \phi^{\beta}$$

where κ is a scaling factor equal to the intercept. In this form the equation represents a power function. S. S. Stevens (1957) proposed that this relationship represents a psychophysical law which states that equal stimulus ratios produce equal sensation ratios.

S. S. Stevens and Galanter (1957) described two general classes of perceptual continua of sensory magnitude, "prothetic" and "metathetic" continua. How continua are classified is determined by how they behave in psychophysical experiments. Prothetic and metathetic continua are thought to be mediated by different physiological processes. Prothetic continua are thought to be associated with changes in sensation magnitude resulting from the addition or subtraction of neural excitation, whereas metathetic continua are thought to be associated with changes in the quality or spatial location of sensation resulting from the substitution of one form of neural excitation for another. Perceptual continua on which subjects make judgments of "how much" (e.g., heaviness, brightness, loudness) belong to the prothetic class of continua. Perceptual continua on which subjects make qualitative judgments of "what kind" and "where" (e.g., pitch, apparent position) belong to the metathetic class of continua. In general, prothetic processes produce perceptual judgment scales in accordance with Stevens' power law, whereas, metathetic processes generally result in judgments that are less orderly (S. S. Stevens, 1957). While loudness sensation may be represented as a power function of stimulus intensity, a scale of pitch sensation assumes a curvilinear form.

Relevant Constructs

Regardless of the psychophysical method employed, quality judgments have normally been used to evaluate speech transmission systems. The signal system, however, is only part of the auditory communication process. Distortion at any point in the transmission or "reception" of a speech signal may influence the signal's auditory perception to some degree. This concept is conveyed by two models. According to information theory (Shannon and Weaver, 1949, 1963),

$$I \text{ (amount of information)} = 2\ t\ w \log (S + N/N)$$

where t is the signal duration, w is the width of the usable frequency range, S is the maximum amplitude of the signal, and N is the minimum discernable intensity difference. Information may be viewed according to what is "received" as well as what is transmitted. Lassman (1964) considered a "noise interference" model in which a hearing aid is the signal transmission system. According to Lassman's model,

$$I = \frac{S/N \text{ environment}}{N_{\text{hearing aid}} + N_{\substack{\text{peripheral}\\ \text{auditory}\\ \text{system}}} + N_{\substack{\text{central}\\ \text{auditory}\\ \text{system}}}}$$

where I is intelligibility, S is signal magnitude, and N is noise in the information theory sense (i.e., anything that increases signal ambiguity). Neither the information theory model nor Lassman's model has been directly related to speech quality judgments, and no claim is made for the validity of either of them. These models are presented only to suggest a relationship between the effects of degradation in the auditory system and in signal transmission systems (e.g., changes in the amount of information passed, degree of intelligibility, and perhaps changes in speech quality).

Apparently, there are no widely accepted theoretical models of speech quality. McGee (1965) discussed two "theories" of speech quality which appear to be insufficiently developed to warrant their appellation. One was a theory of intelligibility based on the articulation index described by French and Steinberg (1947), Fletcher and Galt (1950), and ANSI S3.5-1969. The articulation index involves the construction of a ratio scale from interval data, i.e., percentage scores on articulation tests. In apparent disdain for the intelligibility theory, McGee noted that a perceptual study of speech quality requires a different response by the subject than a written report of what is heard over a speech transmission circuit. The second theory referred to was that of Ochiai and Fukamura (1953, 1956), which McGee called a "vocalic voice" theory. Based upon a thorough analysis of five Japanese vowels, this theory was said to consider a naturalness factor ("vocalic quality") as well as an intelligibility factor ("phonal quality"). The perception of articulation quality was thought to be more directly associated with select portions of the speech spectrum, while the perception of naturalness quality was thought to have a more subtle association with the entire spectrum.

Statement of the Problem

There has been very little systematic research on how the overall quality of complex signals is processed by listeners with normal hearing and listeners with sensorineural hearing loss. The research by audiologists has primarily involved hearing aid transduced connected discourse. Although several studies have demonstrated interest in the use of speech quality judgments as a clinical tool, there are only two systematic studies, Punch and Ciechanowski (1977) and Chial and Daniel (1977), which included sensorineurally impaired listeners. Only one of these two, Chial and Daniel (1977), investigated an alternative psychophysical method to paired-comparisons. This study used ordinal scaling. No study has employed Stevensonian ratio scaling to examine speech quality magnitude estimates as a function of subject groups, types of degradation, and degree of degradation.

Purpose

This study was designed to examine psychophysical functions obtained on listeners with normal hearing and listeners with sensorineurally impaired hearing. Speech quality magnitude estimates (SQMEs) were obtained on the two groups of listeners as a function of three types and seven degrees of signal degradation. The log geometric means of two within-session SQMEs were plotted as a function of log degree of degradation for the two listener groups under the three degradation types. That is, the data were plotted in a log-log space as suggested by S. S. Stevens (1957). These log-log functions were examined to answer the following questions:

1. Is there a statistically significant trend for the log geometric mean SQMEs to be influenced by changes in degree of degradation for each listener group under each degradation type?

2. If a statistically significant trend is present, what is the lowest order equation required to provide a satisfactory (i.e., statistically significant) fit to the data obtained for each listener group under each degradation type?

3. Is there a statistically significant difference among the slopes of the log-log functions as a function of:

    a. listener group (i.e., hearing acuity),

b. degradation type, or

c. the interaction of listener group and degradation type?

CHAPTER II

METHOD

Subjects

Twelve normal hearing listeners (Group 1) and 12 sensorineurally impaired hearing listeners (Group 2) participated in the study. Ages of the Group 1 listeners ranged from 22 to 28 years with a mean of 23.58 years; ages of the Group 2 listeners ranged from 18 to 49 years with a mean of 32.58 years. For Group 1 the average threshold for 500 and 1000 Hz ranged from 0 to 5 dB HTL (re: ANSI S3.6-1969) with a mean of 0.42 dB; for Group 2 the two-frequency average threshold (Fletcher, 1950) for the better ear ranged from 15 to 50 dB HTL with a mean of 33.50 dB. In general, the Group 2 listeners had hearing losses which were more severe for frequencies greater than 1000 Hz. Appendix A shows the age and audiometric data for groups and individual subjects.

Normal Hearing Listeners

The listeners in Group 1 demonstrated normal hearing by:

(1) passing a pure tone screening test at hearing levels of 15 dB (re: ANSI S3.5-1969) for 250, 2000, 4000, and 6000 Hz;

(2) exhibiting hearing threshold levels better than 15 dB (re: ANSI S3.6-1969) at 500 and 1000 Hz;

(3) exhibiting normally shaped tympanograms indicating normal middle ear pressure (Jerger, 1970);

(4) exhibiting acoustic reflex thresholds at hearing levels greater than 60 dB and less than 110 dB (re: ANSI S3.6-1969) at 500, 1000, and 2000 Hz;

(5) exhibiting the ability to sustain stable acoustic reflexes for

10 seconds at 500 and 1000 Hz (Anderson, Barr, and Wedenberg, 1970);

(6) reporting no history of otologic surgery, family hearing loss, recent upper respiratory problems, vertigo, tinnitus, or hearing loss.

In addition, a speech discrimination score of 90% or better was required in the test ear on a commercial version (Auditec of St. Louis) of the Northwestern University Auditory Test Number 6 (NU Auditory Test No. 6). The test was presented at 40 dB above the two-tone average threshold (Fletcher, 1950). Wilson, Coley, Haenel, and Browning (1976) concluded that for clinical purposes the Auditec and original Northwestern versions of the test were equivalent.

Listeners with Sensorineural Hearing Loss

The impaired listeners (Group 2) were initially selected from the clinic records of a hearing and speech clinic. They exhibited sensorineural hearing loss bilaterally, as indicated by:

(1) hearing threshold levels greater than 15 dB at two or more test frequencies (250, 500, 1000, 2000, 4000, and 6000 Hz) as measured by pure tone air conduction tests (Carhart and Jerger, 1959);

(2) normally shaped tympanograms with normal middle ear pressure (Jerger, 1970);

(3) the absence of market tone decay on a tone decay test (Olsen and Noffsinger, 1974) at 1000 and 4000 Hz;

(4) no history of otologic surgery, recent upper respiratory problems, vertigo, or active tinnitus at the time of testing.

In addition, the speech discrimination score in the ear with the better two-tone average threshold was within the 60-90% range on the NU Auditory Test No. 6 presented at a sensation level of 40 dB above the two-frequency average threshold (Fletcher, 1950).

Stimuli

Speech Stimuli

The speech stimuli represented different types of signal degradation and different degrees of degradation within each type.

Talker. Because female talkers seem to accentuate quality differences (Punch and Ciechanowski, 1977), the stimuli were spoken by a female speaker of General American dialect. The following instructions were given to the talker:

1. Use a normal inflectional pattern that is not flat or monotonous.
2. Use normal linguistic emphasis.
3. Do not follow a rhythmic pattern.
4. Use a normal speaking rate.
5. Speak within your normal fundamental frequency range.
6. Peak the VU meter at zero.
7. Always speak from the same physical position (i.e., sitting or standing).
8. Maintain a constant mic-mouth distance of one hand-span.

Stimulus materials. Following the example of other studies involving listening tasks (e.g., Zerlin, 1962; Powers and Speaks, 1973; Chial, 1973; Speaks and Trooien, 1974; Gray and Speaks, 1977; Chial and Daniel, 1977; Punch and Ciechanowski, 1977), this study used orally presented passages of continuous discourse exerpted from reading materials as stimuli. The passages were chosen from three consecutive chapters of a junior high school history test (Wilder, Ludhum, and Brown, 1954). It was felt that this would decrease the diversity in writing style and increase the probability that the passages would be roughly equivalent in readability on a fairly easy level. In addition to the procedure followed by previous researchers, the passages were evaluated by Fang's (1966, 1967) "Easy Listening Formula" (ELF). An "average" ELF score equals the number of syllables above one per word in a sentence divided by the number of sentences. Citing differences between materials prepared for listening (newscasts) and materials prepared for reading (newspapers), Fang concluded that an "average" ELF score below 12 is considered desirable for mass listenability. In the present investigation the ELF was simply used as an index by which the variance in listening ease might be reduced. No claim is made for the reliability or validity of the ELF.

Initially, the ELF was applied to a pool of 93 different passages. Since there is not universal agreement on the definition of a syllable, 17 of the passages were selected for evaluation by three independent raters. Selections were made on the basis of similar ELF scores and a spoken duration of about 10 seconds. At least two of the three independent raters assigned the same score to 14 (82%) of the 17 passages. Of these 14, six were selected to serve as stimuli. Two of the three raters gave an ELF score of 8.5 to five of the six passages and an ELF score of 9.0 to the remaining passage. The absolute deviation among the three raters was 0.5 for the six passages selected. Appendix B is a transcript of the six stimulus passages.

Types of signal degradation. Three degradation types for which articulation and "immediate" intelligibility data are available were selected: (1) low-pass filtering (French and Steinberg, 1947; Chial, 1973), (2) high-pass filtering (French and Steinberg, 1947; Chial, 1973), and (3) linear rectification (Licklider, 1946; Chial, 1973).

Degrees of degradation. Marks (1974) noted that the choice of stimulus range and stimulus spacing over a region of the stimulus scale should be based on the range and spacing of sensory magnitudes rather than stimulus magnitudes. Insofar as possible, sensory data were employed to determine degrees of degradation. In the magnitude estimation functions reported by S. S. Stevens (1975), five to eight points were usually plotted. Thus, it was felt that seven degradation levels should provide sufficient data to permit estimation of appropriate functions. For each degradation type, one of the seven degradation levels was a "no" degradation condition.

French and Steinberg (1947) reported syllable articulation data which were obtained at constant intensities and plotted as a function of the cutoff frequency of a low-pass and a high-pass filter. In the present study an all-pass or "no" degradation condition was arbitrarily assigned an articulation score of 100% for the low-pass and high-pass "filtered" speech. Thereafter, cutoff frequencies were interpolated from French and Steinberg's data for syllable articulation decrements of 15% (i.e., 85, 70, 55, 40, 25, and 10%). The cutoff frequencies for low-pass filtering were 3000, 2000, 1700, 1350, 1000, and 600 Hz. The cutoff frequencies for high-pass filtering were 1400, 1850, 2300, 2800, 3000, and 3500 Hz. It was felt that the choice of cutoff frequencies which produce approximately equal decrements in articulation scores might lead to conclusions regarding the intelligibility "theory" of speech quality referred to by McGee (1965).

Degrees of linear rectification were chosen on the basis of percent total harmonic distortion (THD) produced by signal rectification. Starting with less than 1% THD (the "no" degradation condition), seven magnitudes of THD were selected: 1%, 10%, 20%, 30%, 40%, 50%, and 60%. Both Licklider (1946) and Chial (1973) included a half-wave rectification condition, which Chial determined to represent approximately 40% THD. Thus, the selected THD values were directly related to previous data by the 40% THD value.

Master recording of undegraded stimuli. The following apparatus was used to produce a master recording of the undegraded speech passages. A microphone (Electrovoice model RE-15) and a VU meter were located in the double-walled test room of a sound suite (IAC 1200 Series). The output of the microphone was passed through the wall to one channel of a VU meter bridge (Teac MB-20) and associated audio mixer (Teac 2). The output of the mixer was routed to a two-track reel-to-reel magnetic tape recorder (Ampex AG-500). The output of the tape recorder was passed through another channel of the audio mixer and meter bridge to a remote VU meter in front of the speaker. The six passages selected as stimuli were spoken by a female speaker in the test room and recorded on a master tape at a speed of 7.5 inches per second. The same speaker also produced stimulus labels and headings (e.g., "Trial 1", "Item 3", etc.).

Submaster recordings of filtered stimuli. The master signal from the Ampex recorder was passed through a filter set (Krohn-Hite 3550) to a cassette recorder (Nakamichi 700 II) where it was re-recorded in varying degrees of degradation by high- or low-pass filtering. Prior to recording a speech sample of any type and degree of degradation, however, the apparatus shown in Figure 1 was used to calibrate and verify the filter cutoff frequencies. The output of a swept sine generator (Bruel and Kjaer 1024) was passed through an attenuator set (Hewlett-Packard 350D) to the filter set. The input of the filter set was monitored by a frequency counter (Heath/Schlumberger SM 4100). The output of the filter set was monitored on an RMS voltmeter (Bruel and Kjaer 2607) and recorded by a graphic level recorder (Bruel and Kjaer 2305). Response curves are depicted in Figure 2 for the low cutoff settings and in Figure 3 for the high cutoff settings.

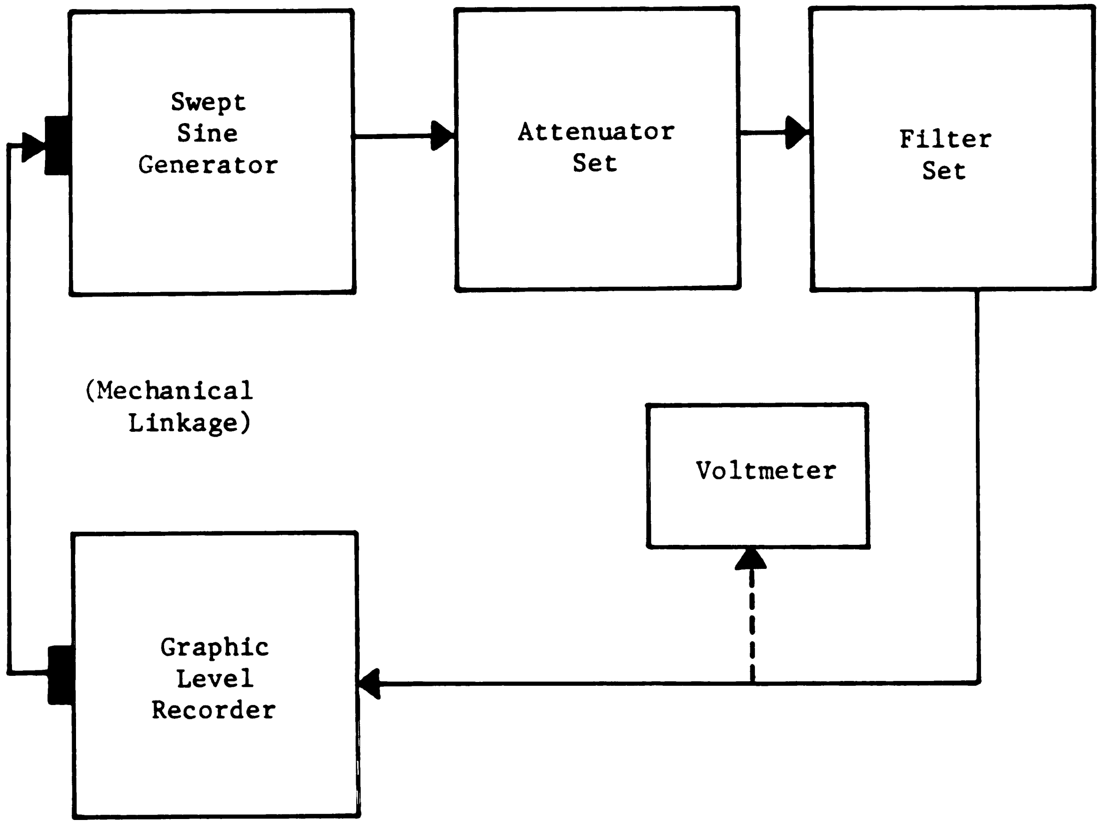

Figure 1. Apparatus used to calibrate and verify filter cutoff frequencies.

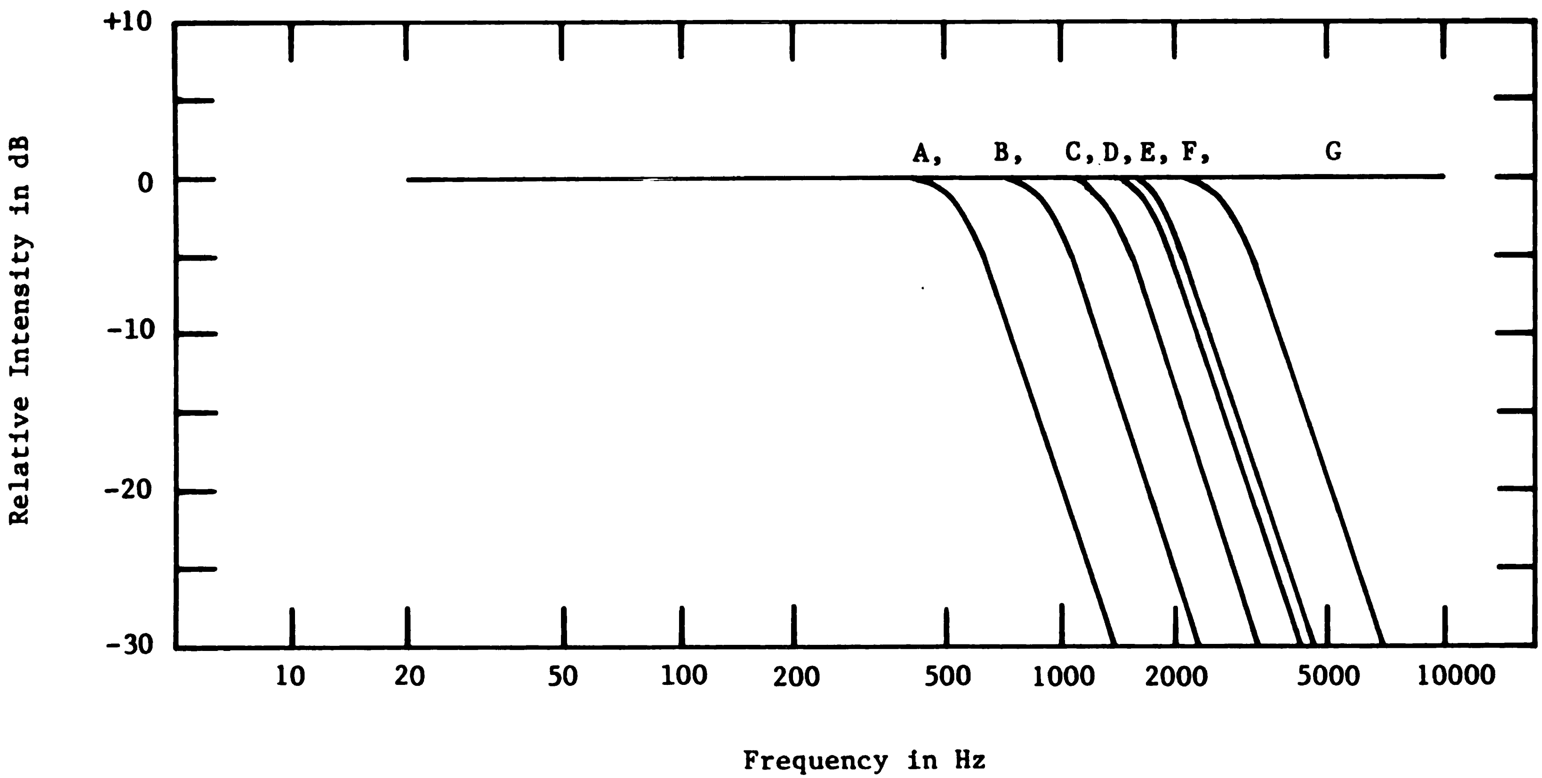

Figure 2. Frequency response curves for the Krohn-Hite Filter set at selected low-pass cutoff frequencies: (A) 600 Hz, (B) 1000 Hz, (C) 1350 Hz, (D) 1700 Hz, (E) 2000 Hz, (F) 3000 Hz, and (G) "none".

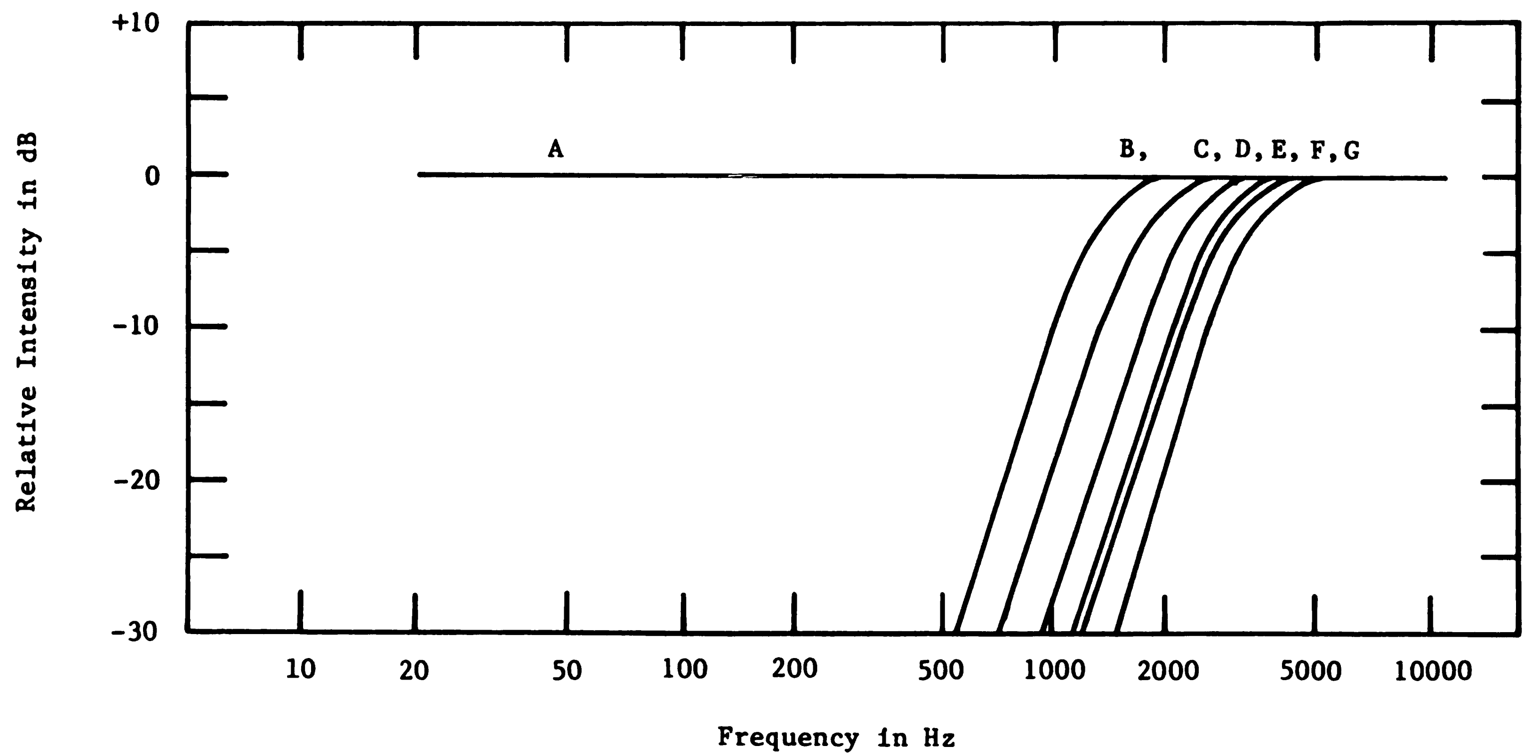

Figure 3. Frequency response curves for the Krohn-Hite Filter set at selected high-pass cutoff frequencies: (A) "none", (B) 1400 Hz, (C) 1850 Hz, (D) 2300 Hz, (E) 2800 Hz, (F) 3000 Hz, and (G) 3500 Hz.

Submaster recordings of rectified stimuli. The master signal from the Ampex recorder was passed through the variable rectified portion of a custom-built CD-1 speech distortion instrument described by Chial (1973). The output of the variable rectifier was passed to the Nakamichi cassette recorder. Prior to recording speech samples, however, the apparatus shown in Figure 4 was used to calibrate the rectifier for the desired percent total harmonic distortion values (% THD). A 1 kHz signal from a sine generator (Bruel and Kjaer 1024) was fed through an attenuator (Hewlett-Packard 350D) to the variable rectifier. The output of the sine generator was monitored by a frequency counter (Heath/Schlumberger SM 4100) and by one channel of a dual beam bistable storage oscilloscope (Tektronix 5113). The output of the variable rectifier was monitored by the other channel of the oscilloscope, as it passed to a frequency analyzer (Bruel and Kjaer 2107). Thus, the undistorted and the distorted waveforms could be monitored simultaneously.

The output of the variable rectifier was adjusted by approximation until the desired THD values were measured on the frequency analyzer. In addition to reading direct THD values from the frequency analyzer, the amplitudes of the first through the tenth harmonics of the 1 kHz driving signal were also obtained. This allowed computation of THD according to the standard formula suggested by ANSI S3.3-1960,

$$\%\ \mathrm{THD} = 100 \cdot \sqrt{\frac{A_2^2 + A_3^2 + \dots A_n^2}{A_1^2 + A_2^2 + \dots A_n^2}}.$$

The results of the harmonic distortion measurements (Table 1) showed good agreement between the computed THD and the direct THD.

Once a desired THD value was obtained, the wave forms monitored on

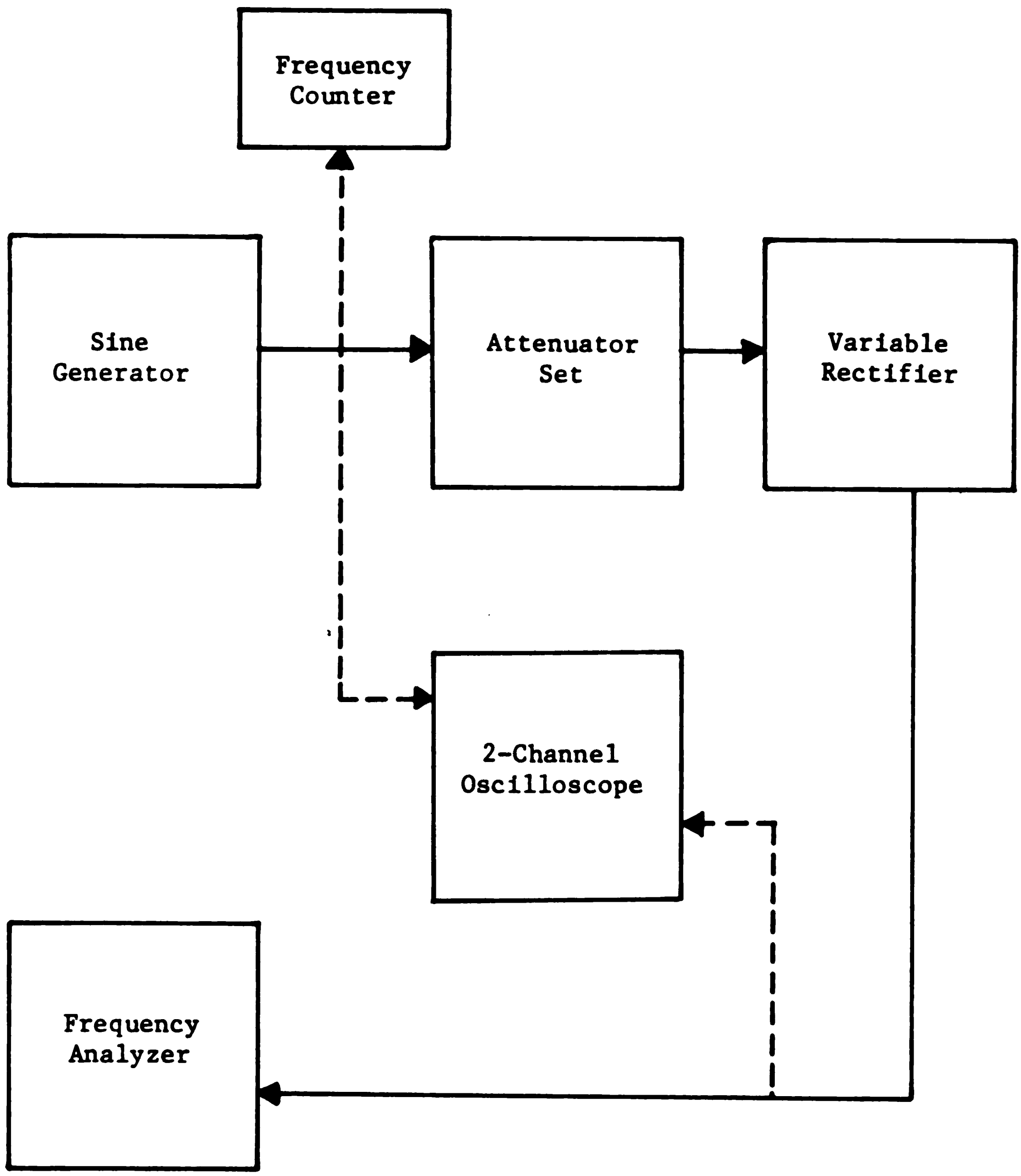

Figure 4. Apparatus used to calibrate the variable rectifier.

Table 1. Harmonic amplitude measurements in millivolts and computed percentages (%) of total harmonic distortion (THD) in the output of the variable rectifier at seven different settings. The variable rectifier was adjusted to achieve seven degrees of THD as measured by a direct method for a 1 kHz driving signal.

	Percent THD as Measured by the Direct Method						
Harmonic	"None" (0.63)	10	20	30	40	50	60
1	820.00	800.00	800.00	740.00	740.00	690.00	620.00
2	6.40	75.00	160.00	225.00	310.00	390.00	460.00
3	2.90	3.40	4.90	6.70	8.50	10.50	12.50
4	2.00	15.50	32.00	46.00	63.00	78.00	92.00
5	1.80	2.40	3.20	4.50	5.50	7.00	8.50
6	1.60	6.60	13.50	19.00	26.50	33.00	39.00
7	1.50	1.90	2.80	4.00	5.00	6.20	7.50
8	1.45	4.00	7.80	11.00	15.50	15.50	22.50
9	1.40	1.70	2.60	3.50	4.50	5.60	6.70
10	1.35	2.70	5.10	7.10	10.00	12.00	14.50
Computed % THD*	1.00	9.60	20.10	29.80	39.50	50.14	60.56

*Computed % THD $= 100 \cdot \sqrt{\dfrac{A_2^2 + A_3^2 + \cdots A_{10}^2}{A_1^2 + A_2^2 + \cdots A_{10}^2}}$

the oscilloscope were photographed for future use. These permanent calibration waveforms (Figure 5) made it possible to reproduce the desired THD values. Subsequently, the variable rectifier was simply adjusted to achieve the appropriate calibration waveform, and the percent THD was verified by the direct method on the frequency analyzer.

Summary of submaster recordings. The master recording was used to produce nine cassette submasters which are described in Appendix C. Each of the first seven submasters consisted of comparison passages A, B, and C under one of the seven degrees of degradation for each degradation type. The eighth submaster consisted of standard comparisons A, B, and C under a single degree of degradation for each degradation type. The ninth submaster consisted of labels (i.e., "Trial 1", Trial 2", "Item 1", etc.) which were dubbed directly from the master recording.

Computer generated tapes. The cassette submaster recordings were used with the computer system shown in Figure 6 to generate reel-to-reel recordings. The cassette recorder fed the signal from the submaster tapes into a 3 Rivers Computer Corporation analog-to-digital converter (ADC). The digitized signal from the ADC was passed to a digital computer (Digital Equipment Corporation pdp 11/40) which was interfaced with two disk drives (Digital Equipment Corporation RK05), a teletype terminal (Digital Equipment Corporation, decwriter II), and a video monitor. The digitized signals were processed by the computer to generate different random orders of stimuli and to control time intervals and stimulus output levels. The digital computer was patched to a 3 Rivers Computer Corporation digital-to-analog converter (DAC) and then to a reel-to-reel recorder.

The computer system was controlled by three interactive programs: GLADC (Appendix D), RPLAY (Appendix E), and GSCALE (Appendix F). GLADC

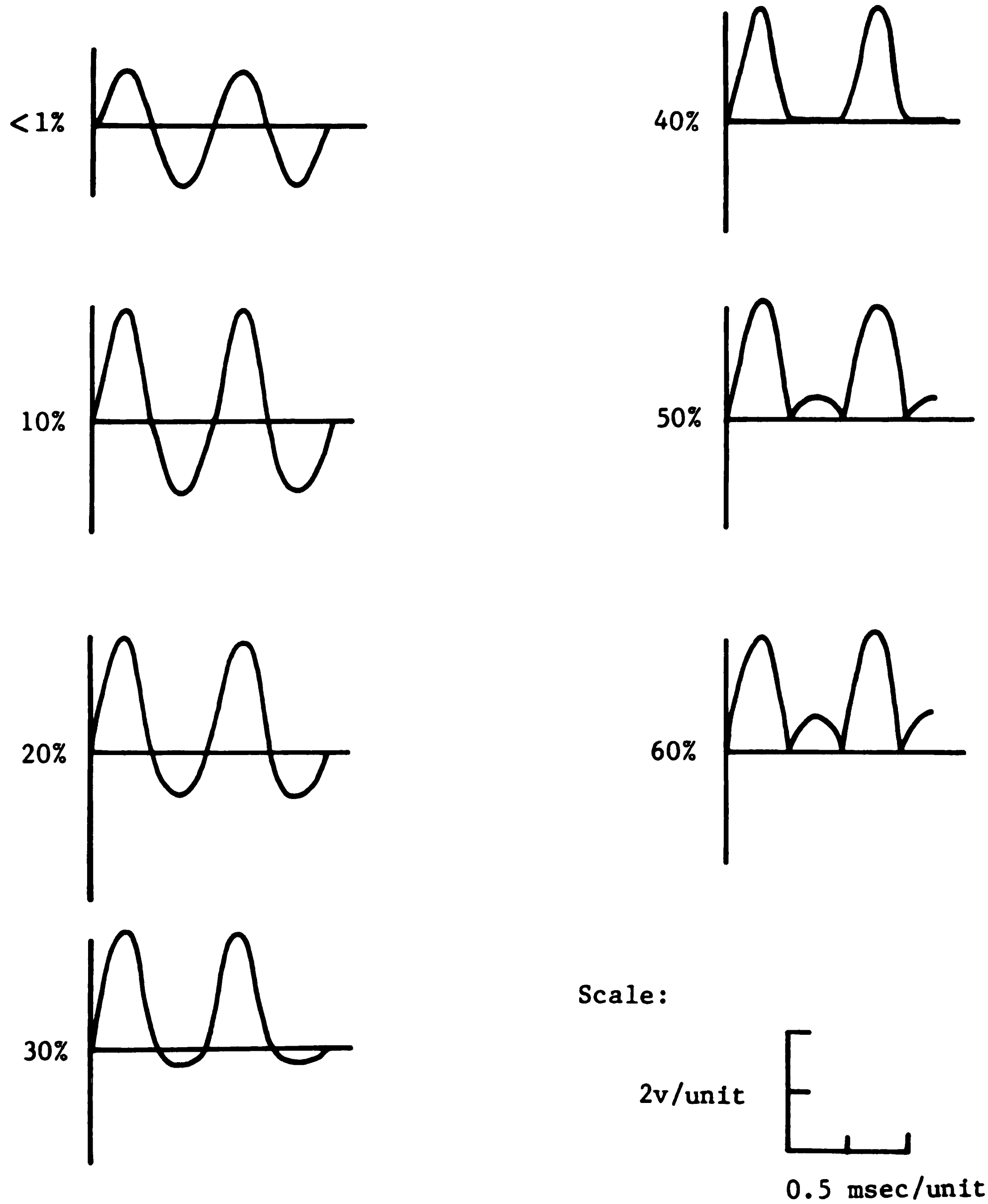

Figure 5. Calibration waveforms used to obtain desired percentages of total harmonic distortion.

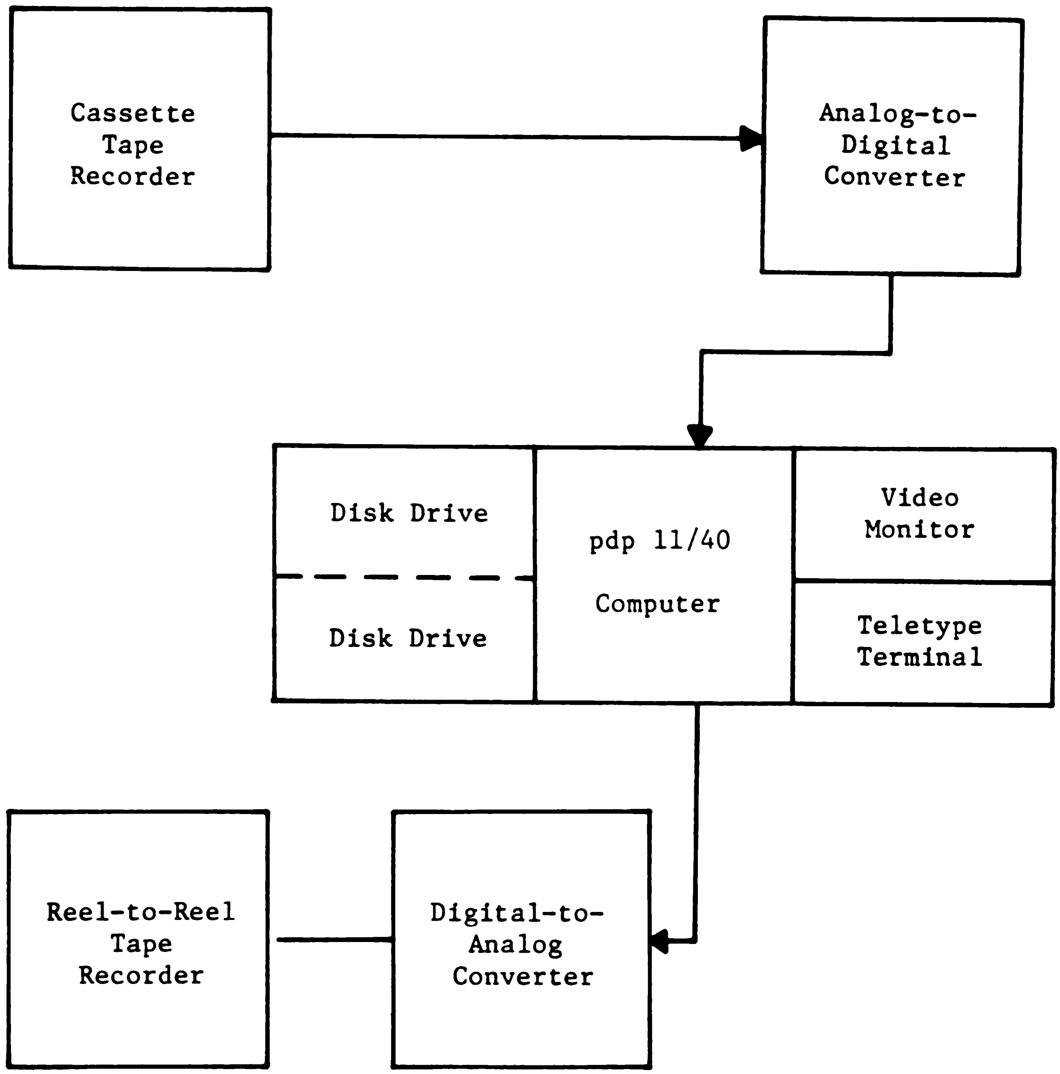

Figure 6. Computer system used to generate duplicate stimuli and play them back in random orders.

was used for analog-to-digital conversions, that is to create sound files, while RPLAY was used to convert the digital sound files to analog form and to play them in random order. A 10 kHz sample rate was used in the conversion programs. GSCALE was used to make any necessary adjustments in playback level required to achieve zero VU on the Ampex recorder.

Permanent sound files were created for a 1 kHz calibration tone and for voiced headings and labels (e.g., "Trial 1", "Trial 2", Trial 3", "standard", "item", and spoken digits "1" through "7". Trial headings were used to announce the beginning of each series of seven stimulus pairs. The word "standard" followed each trial heading to introduce the standard passage for the upcoming series of stimuli. The word "item" always preceded a number (e.g., 1-7) used to designate a stimulus pair. Each pair consisted of the standard stimulus followed by a comparison stimulus. Temporary sound files were created for the standard passage and for each of the seven degradation levels of a given comparison passage under a degradation type.

Nine reel-to-reel recordings (Appendix C) were produced by the computer system. Each of these interim tapes represented a single degradation type and consisted of six different random orders, or trials, of seven stimulus pairs. In each of these pairs, the standard passage always represented the middle or fourth degradation level; the comparison passage represented one of the seven degradation levels.

Final test tapes. Six final tapes were prepared by splicing timing (leader) tape and trial segments from the tapes generated by the computer. Appendix C shows the presentation orders for the three degradation types, six spoken passages, and seven degrees of degradation of the comparison passages. Each final test tape consisted of (1) a level calibration tone, (2) spoken instructions to the subject, (3) practice materials including one seven-item trial for each of three degradation types, and (4) experimental materials including two seven-item trials for each of the three degradation types.

Each trial lasted approximately 3.5 to 4.0 minutes. Approximately 15 seconds were required for the trial heading, standard heading, and standard passage. Approximately 3.5 minutes were required for the seven stimulus items (i.e., about 30 seconds for each of the 7 items). The standard and comparison passages lasted approximately 10 seconds each and were followed by a 5-second response interval.

Effects of apparatus. The final auditory stimuli received by each listener were subject to filtering imposed by all instruments used to generate and present the stimuli. The apparatus used in stimulus generation was described earlier. A block diagram of the listening apparatus is shown in Figure 7. In the control room the output of a reel-to-reel tape recorder (Ampex AG-500) was routed to the tape input of a two-channel speech audiometer (Grason-Stadler 162), then through the wall to two pairs of TDH-49 earphones located in the test room. One earphone of each pair was a "dummy".

Appendix G describes the measurement of frequency response curves of instruments used to generate and present auditory stimuli. Composites of these response curves (Figure 8) show that (a) the highest of the low frequency cutoffs (i.e., the 3-db down points) was imposed at 50 Hz by the TDH-49 earphones and (b) the lowest of the high frequency cutoffs was imposed at 4400 Hz by the computer system. The high frequency cutoff was due primarily to the use of an ADC sampling rate of 10 kHz. A higher sampling rate would have been desirable but was precluded by computer system limitations. These response curves show that all "undegraded" and linearly rectified speech stimuli had a low frequency cutoff of 50 Hz and

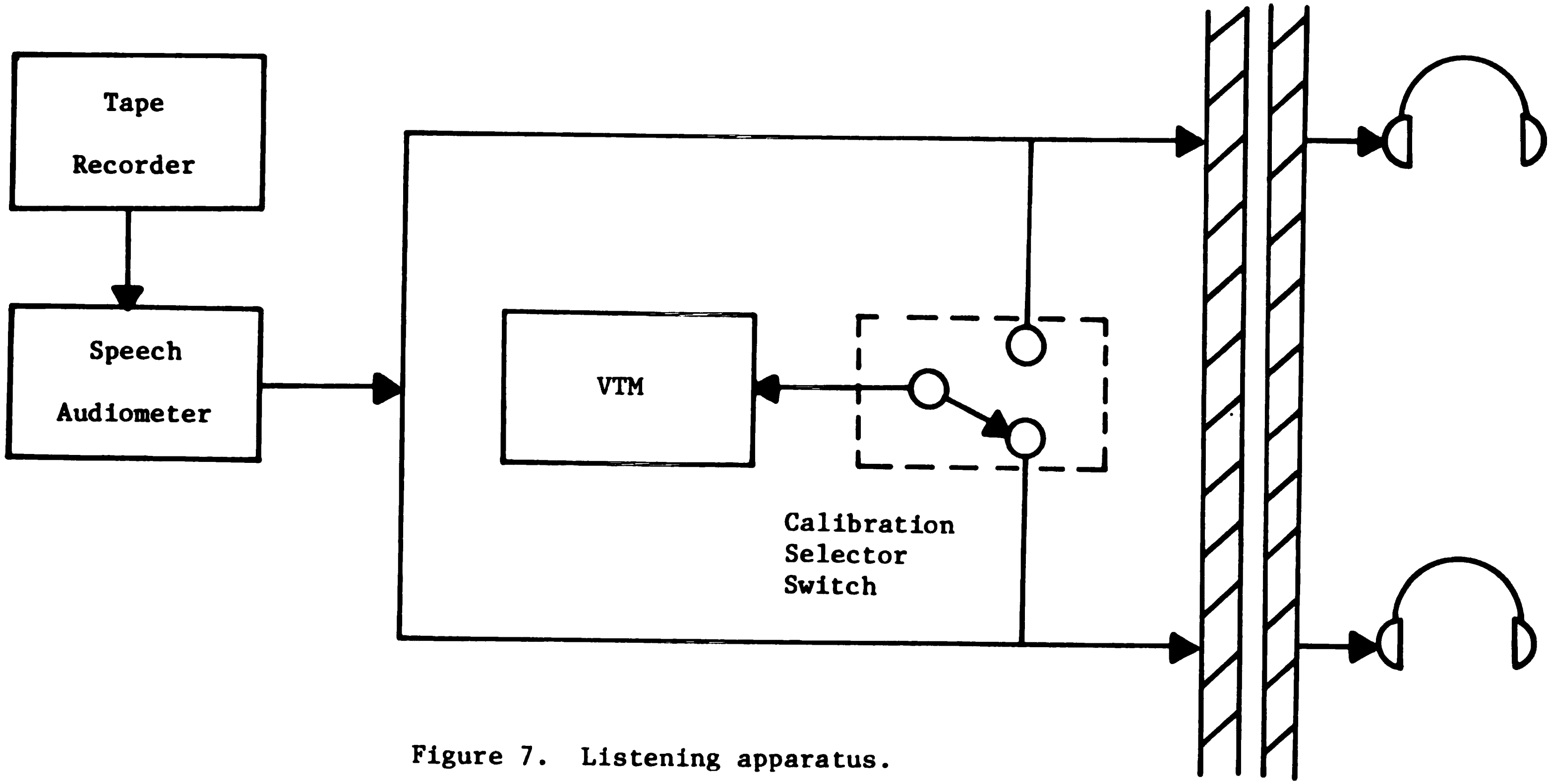

Figure 7. Listening apparatus.

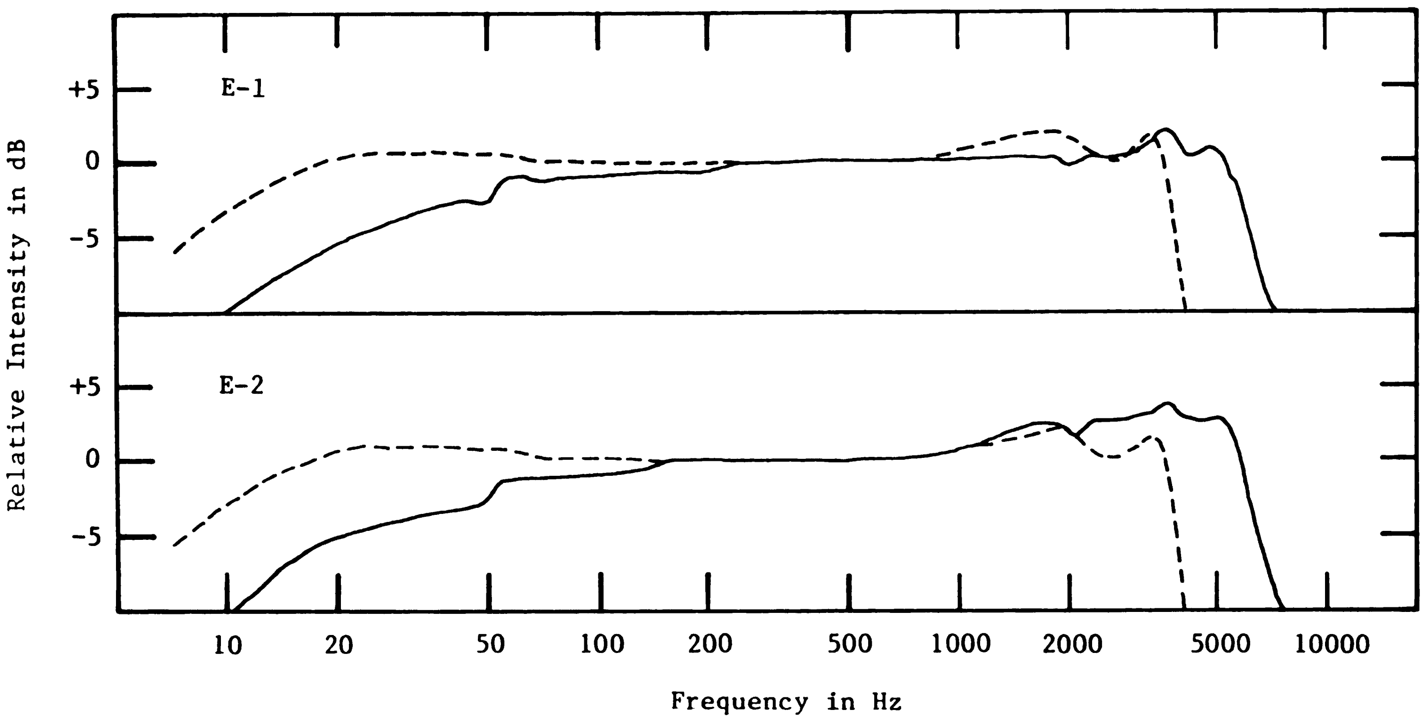

Figure 8. Composite frequency response curves for the computer system (dashed lines) and the Grason-Stadler 162 speech audiometer with TDH-49 earphones (solid lines for E-1 and E-2).

a high frequency cutoff of 4400 Hz. These cutoff frequencies, together with those imposed by the low-pass and high-pass degradation conditions, effectively determined bandwidths for all the filtered stimuli.

Visual Training and Screening Stimuli

Naive and unpracticed subjects are apparently capable of yielding reliable and consistent results on magnitude estimation tasks (S. S. Stevens and Poulton, 1956; J. C. Stevens and Tulving, 1957). However, S. S. Stevens (1975) noted that it is sometimes helpful to initiate the new observer with an easy experiment such as the judgment of apparent line length or circle size. Also, it seemed important in this study to eliminate subjects who showed great difficulty with the method of magnitude estimation. The goal was to eliminate those who had difficulty with the method, not those who had difficulty judging speech quality. Thus, a visual screening and training task seemed appropriate.

Three sets of visual stimuli were used for the screening and training task. Each set represented a trial comprised of seven pairs of stimuli produced on 35-mm slides. Each stimulus pair consisted of two white geometric forms located side by side on a blue background. An example of how the stimuli appeared on slides is shown in Figure 9. Each slide had an identifying number centered beneath the two stimuli. The left member of the pair, the standard stimulus, was located below an "S" for standard. The right member of the pair, the comparison stimulus, was located below a "C" for comparison. The middle-sized comparison shape was the same size as the standard. The first trial consisted of squares of different areas. The second trial consisted of circles of different areas. The third trial consisted of a different ordering of the same size circles used in the second set.

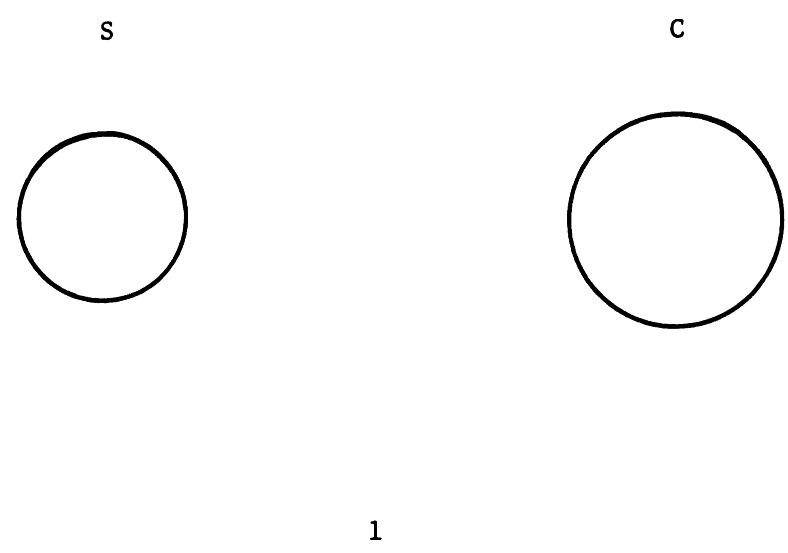

Figure 9. Example of a stimulus slide used in the visual training and screening task.

Procedures

Experimental procedures are summarized in Figure 10. Appendix H shows the protocol followed by the experimenter in running each subject.

Audiometric Screening

Each subject was asked to sign an informed consent release form (Appendix I) and to verify case history data taken earlier by telephone. Subjects were then evaluated to determine their qualifications for the study and to provide reference thresholds for the experimental tasks. Audiological test results and subject history information were recorded on a form devised for that purpose (Appendix J).

With the exception of the speech discrimination tests, all hearing tests were administered to each ear. Normal hearing listeners received a pure tone screening test; hearing impaired listeners received a threshold test (Carhart and Jerger, 1959). Normal listeners were tested for reflex decay at a sensation level of 10 dB (re: acoustic reflex threshold) at 500 and 1000 Hz; hearing impaired listeners received an audiometric tone decay test (Olsen and Noffsinger, 1974) at 1000 and 4000 Hz. Tympanograms were plotted manually for all subjects, taking compliance measurements at pressure increments of 100 mm H_2O.

For each subject, the "better" ear was designated as that ear which produced the lowest two-frequency pure tone average threshold. In the absence of interaural difference, the right ear was arbitrarily selected as the test ear. A speech discrimination score was obtained for the test ear by presenting the NU Auditory Test No. 6 at 40 dB sensation level (re: the two-tone average threshold).

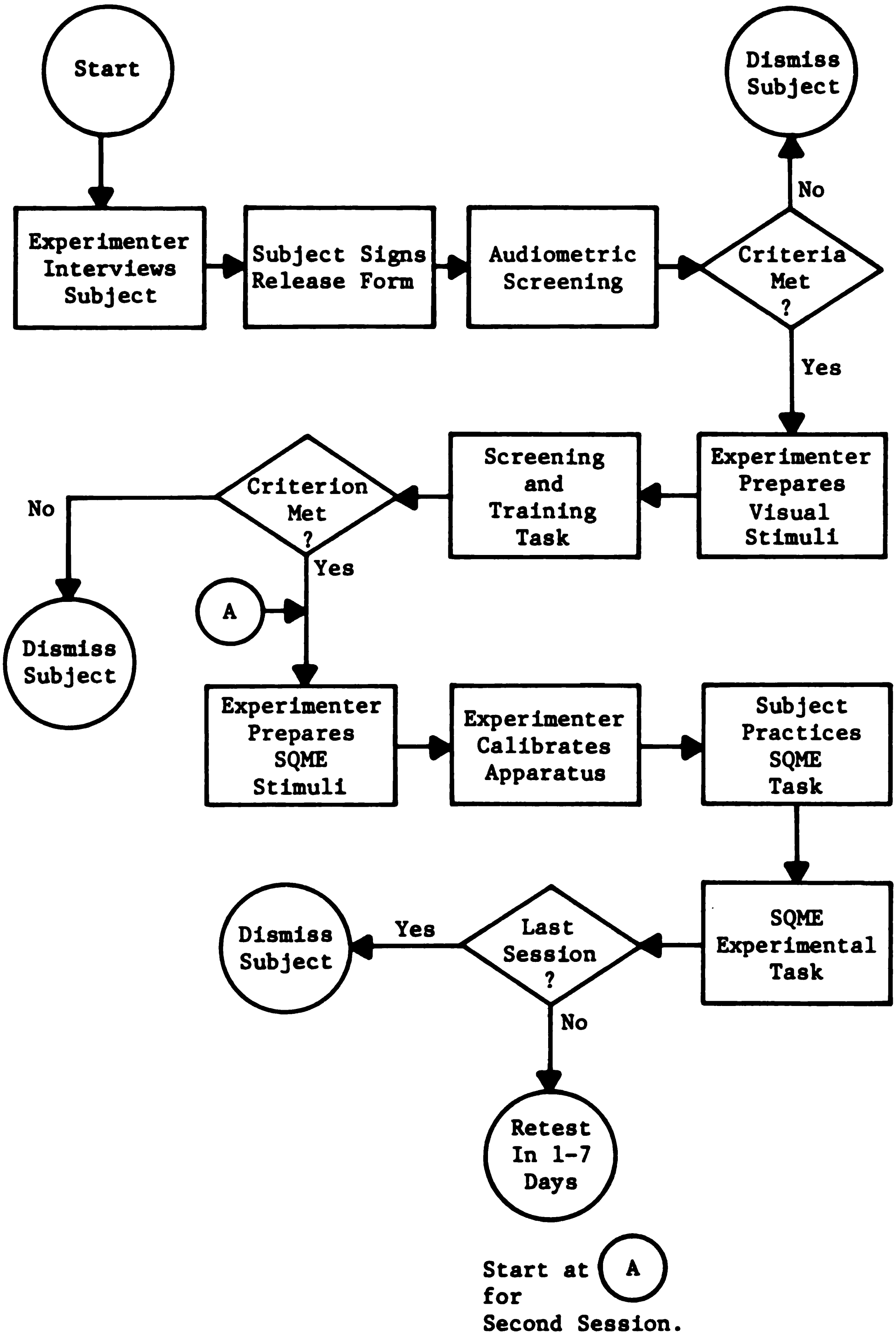

Figure 10. Procedural flowchart.

Visual Magnitude Estimation Training and Screening

Following the audiometric testing, a 15-minute program of audio-taped instructions and stimulus slides was used to teach the subject how to perform a magnitude estimation task and to assess the consistency of the subject's performance in the task. The script of the audio-visual training task is given in Appendix K.

The training and screening program was conducted individually in one room of a double-walled sound suite (IAC 1200 series). The audio signal was presented through earphones at a comfortable loudness level. Each subject estimated the apparent magnitude of various squares and circles which were projected on a rear-screen (9" x 9") slide viewer equipped with a synchronized tape player (Singer Caramate II SP). Instructions to the subject and a sample response sheet are given in Appendix L. The visual training task was intentionally similar in form and procedure to the auditory tasks.

Following a practice trial with randomly ordered squares, each subject made magnitude estimates of circle size in two subsequent trials which represented different randomizations of the same stimuli. Each subject retained for the experiment was required to produce a correlation (r) between estimates for the second and third trials that was equal to or greater than 0.90. Subjects also were retained only if they assigned the same estimate to a standard stimulus and its equivalent comparison stimulus. A pilot study (Appendix M) of a separate group of subjects who met these criteria yielded results similar to those reported by S. S. Stevens (1975).

Listening Tasks

The listening tasks consisted of SQME training and the SQME experiment. Listening tasks followed the visual magnitude estimation training and screening procedures and were administered to the subjects in pairs. Two subjects were located in back-to-back writing desks located in the test room. Stimuli were presented at 40 dB above the two-frequency average threshold (Fletcher, 1950).

Calibration of listening apparatus. The listening apparatus (Figure 7, p. 34) enabled the same signal to be delivered at different intensities to each of two subjects. A calibration selector switch allowed the examiner to monitor the taped calibration tone routed to either experimental earphone. The signal level to the earphones was checked in this manner prior to each experimental session.

In addition to the within-session calibration, the tape recorder, speech audiometer, and earphones were checked before and after the investigation. On both occasions, the system performed within the tolerances specified by ANSI S3.6-1969.

SQME training. Instructions for the listening tasks were given orally (via tape recording) and in writing. The written instructions and samples of the accompanying response sheets are shown in Appendix N. Subjects were instructed to assign any numerical value which seemed appropriate to the standard stimulus for each trial and to assign a related numerical value (i.e., a magnitude estimate of speech quality) to each comparison stimulus. In other words, the training task, like the experiment itself, employed a free modulus paradigm.

The instructions were followed by 21 practice items (seven degrees of each of three types of degradation). The order of degradation types was counterbalanced across pairs of subjects. The same standard passage and the same comparison passage were used for all practice trials, but the order of comparison degradation levels varied randomly within each trial so that a given order was never repeated.

The SQME training lasted approximately 15 minutes and was followed by a 5-minute break.

SQME experiment. Practice stimuli were followed by experimental stimuli on the same reel. The experimental tasks were conducted in the same manner as the practice task. Subjects performed speech quality magnitude estimates on 42 items (two seven-item trials of each of the three degradation types). The order of degradation types was counterbalanced across pairs of subjects, and the order of stimulus passages was counterbalanced across trials. Again, the order of degradation levels varied randomly within each trial so that a given order was never repeated.

The total stimulus time was approximately 21 minutes (3.5 minutes for each of 6 stimulus sets). Subjects received a two-minute break after every two trials, that is, after listening for about 9 minutes. In other words, the subjects completed their judgments on all samples of a single degradation type prior to taking a break.

Second Listening Session

Eight listeners were randomly selected to participate in a second listening session. Four normal hearing listeners and four hearing impaired listeners returned within one to seven days to repeat the training and the SQME experiment.

Summary

A summary of the experimental events and their time requirements is shown in Table 2. The primary screening, training, and listening session lasted approximately 50 minutes.

Table 2. Summary of events and their time requirements.

Event	Time Required (Minutes)
Audiological Screening	
History (taken by phone)	0
Pure tone air conduction testing	10
Tone decay testing	6
Discrimination testing in test ear	6
Impedance testing	10
Break	5
Visual Magnitude Estimation Training and Screening	12
*Speech Quality Magnitude Estimation Training	15
Break	5
Speech Quality Magnitude Estimation Experiment	
Stimulus sets 1 and 2	9
Break	2
Stimulus sets 3 and 4	9
Break	2
Stimulus sets 5 and 6	9
All events	100 minutes

*Second session begins with retraining for SQME.

CHAPTER III

RESULTS

Introduction

The purpose of this study was to examine speech quality magnitude estimation (SQME) functions of listeners with normal hearing and listeners with sensorineural hearing loss. The SQMEs were obtained as a function of seven degrees of each of three types of degradation.

Twelve normal hearing subjects and 12 hearing-impaired subjects met all selection criteria for the study. Two additional hearing-impaired subjects were eliminated because they did not demonstrate adequate consistency in the visual magnitude estimation task.

Initially, the listeners were trained in the use of ratio scaling to perform magnitude estimations of square and circle size. They were then instructed in magnitude estimation of speech quality and presented with 21 practice stimuli (seven degrees of each of three types of degradation). Practice stimuli were followed by 42 experimental stimuli (seven degrees of each of three types of degradation for two trials). Both the practice and the experimental stimuli consisted of connected speech samples which had been degraded by low-pass filtering, high-pass filtering, and linear rectification.

Four subjects were randomly selected from each group to participate in a second experimental session. The second session was conducted in the same manner as the first, beginning with instruction in speech quality magnitude estimation.

Log geometric mean SQMEs obtained over trials were plotted as a function of log degree of degradation. The log-log data were examined to answer the following experimental questions:

1. Is there a statistically significant trend for log geometric mean SQMEs to be influenced by changes in degree of degradation for each listener group under each degradation type?

2. If a statistically significant trend is present, what is the lowest order equation required to provide a satisfactory (i.e., statistically significant) fit to the data obtained for each listener group under each degradation type?

3. Is there a statistically significant difference among the slopes of the log-log functions as a function of:

   a. listener group (i.e., hearing acuity),

   b. degradation types, or

   c. the interaction of listener group and degradation type?

Data Reduction

The experimental questions deal with (a) the log geometric means of magnitude estimates obtained for two trials and (b) the slopes of the functions relating the log geometric mean estimates to log stimulus values. Thus, it was necessary to reduce individual magnitude estimates to log geometric magnitude estimates and slope terms. This was accomplished through the use of MAGEST, a Fortran IV computer program for analyzing magnitude estimation data (Kerst, 1978). The inputs required by this program include the perceptual magnitude estimates and the values of the stimulus magnitudes.

The next two sections describe operations on the auditory stimulus values and the SQMEs from the listening experiment. The procedures employed with the SQMEs were also followed to reduce data from the visual training and screening task.

Auditory Stimulus Magnitudes

It was desirable to express the stimulus magnitudes in a manner that required their numerical values to change in the same direction relative to changes in degradation across degradation types. Thus, auditory stimulus magnitudes were expressed as frequency bandwidths for the filtered stimuli and as percent undegraded values for the linearly rectified stimuli. (Note that the values of the visual stimulus magnitudes were appropriately expressed in their original form, square inches.)

Table 3 summarizes the low-pass and high-pass cutoff frequencies with their respective bandwidths and the percentages of THD with their respective percent undegraded values. Bandwidths of low-pass filtered stimuli were determined by subtracting the lower frequency limit of the listening apparatus (50 Hz) from each low-pass cutoff frequency. The "no" degradation value was determined by subtracting 50 Hz from the upper frequency limit of the computer system used in stimulus preparation (4400 Hz). Bandwidths of high-pass filtered stimuli were determined by subtracting each high-pass cutoff frequency from 4400 Hz (the upper frequency limit of the computer system used in stimulus preparation). The "no" degradation value was determined just as it was for the low-pass filtered stimuli, that is, by subtracting 50 Hz from 4400 Hz. Values for percent undegraded by linear rectification were determined by subtracting the measured percentages of THD from 100% THD. The "no" degradation level (0.63% THD) was called 1% THD or 99% undegraded.

Numerical values for degrees of degradation were entered into the MAGEST program separately for each degradation type. MAGEST transforms these values by taking the natural log of each. The log stimulus values are retained by the program as independent variables to be used in

Table 3. Degrees of degradation expressed as (1) stimulus cutoff frequencies and their respective bandwidths for the filtered stimuli and (2) percentages of total harmonic distortion (% THD) and their respective percent undegraded values for linearly rectified stimuli. The seven levels of each degradation type are listed from left to right in order of increasing degradation.

<u>Degrees of Degradation by Low-Pass Filtering</u>						
A. Low-pass cutoff frequency (Hz)						
4400*	3000	2000	1700	1350	1000	600
B. Low-pass bandwidth (Hz)						
4350	2950	1950	1650	1300	950	550
<u>Degrees of Degradation by High-pass Filtering</u>						
A. High-pass cutoff frequency (Hz)						
50*	1400	1850	2300	2800	3000	3500
B. High-pass bandwidth (Hz)						
4350	3000	2550	2100	1600	1400	900
<u>Degrees of Degradation by Linear Rectification</u>						
A. Percent (%) THD						
1%*	10%	20%	30%	40%	50%	60%
B. Percent (%) undegraded						
99%	90%	80%	70%	60%	50%	40%

* "no" degradation

subsequent processing of SQMEs obtained from individual subjects.

SQMEs

SQMEs were tabulated from individual response sheets for each subject and each listening condition. They were then coded and logged on punch tape to create data files for subsequent processing by MAGEST. MAGEST computes geometric mean and log geometric mean SQMEs across trials for each subject and for each degree of signal degradation. The geometric mean (G.M.) is defined as

$$G.M. = \sqrt[n]{(X_1)(X_2)\ldots(X_n)}$$

where x_n is the nth score. The geometric mean is the preferred index of central tendency (S. S. Stevens, 1975) because (a) it is consistent with the underlying scale of measurement (ratio) and (b) it is relatively insensitive to the effects of modulus differences across subjects or trials. MAGEST also applies the method of least squares to determine the slope and intercept of the linear equation which relates the log geometric mean SQMEs and log stimulus values. These computations were accomplished both for individual subjects across trials and for listener groups across subjects. Other MAGEST outputs are described by Kerst (1978).

MAGEST was executed several times to produce log geometric mean SQMEs and slopes for individual subjects and for groups of subjects. The program was run separately for all combinations of degradation types, trials, and groups.

Statistical Procedures

In a magnitude estimation task, the values of magnitude estimates or log geometric mean magnitude estimates across trials are directly affected by the numerical value of the standard stimulus, that is, the modulus. The slopes of the linear equations which relate the log magnitude estimates and log geometric mean magnitude estimates to log stimulus values, however, are relatively independent of the modulus chosen. Since all magnitude estimates were made in a modulus-free manner, only within-subject analyses were done on the magnitude estimates and log geometric mean magnitude estimates across trials. Both within- and between-subject analyses were done on slopes. A significance level of 0.05 was used in all statistical tests.

Reliability Procedures for SQME Data

Pearson product-moment correlation coefficients (Linton and Gallo, 1975, pp. 347-352) were computed to assess reliability within and between experimental listening sessions.

Within sessions. Within-session correlation coefficients were obtained for variables underlying the dependent variables in Session 1 and Session 2. Correlation coefficients were obtained between each of the 24 subject's Trial 1 and Trial 2 SQMEs under each degradation type in Session 1. Thus, coefficients between Trial 1 and Trial 2 SQMEs for the two subgroups (n = 4) were available as subsets of the data for Session 1; additional subgroup coefficients had to be determined only for Session 2. Correlation coefficients for groups and degradation types were transformed to Fisher Z scores (Hays, 1963, pp. 680-681), which were summed and divided by the appropriate number of scores to determine mean Z scores. The mean Z scores were then transformed to correlation coefficients to obtain "average" coefficients.

Between sessions. Between-session correlation coefficients were obtained for the dependent variables considered in the experimental questions. The dependent variables were log geometric mean SQMEs across trials and the slopes of the lines relating the log geometric means to log stimulus values. Between-session data were available only for those four subjects in each group who returned to participate in a second listening session. Each subject's seven log geometric mean SQMEs for Session 1 were correlated with those for Session 2 under each of the three types of degradation. The slopes obtained in Session 1 were correlated with those obtained in Session 2 for each listener group (n = 4) under each degradation type. "Average" between-session correlation coefficients were determined by interpolating Z to r transformations from mean Fisher's r to Z transformations.

Analysis Procedures for SQME Data

Perceptions of speech quality were examined first by analyzing the log geometric mean SQMEs as a function of log stimulus values, and second, by analyzing the slopes of the log-log functions.

The log geometric means were examined separately for each of the two listener groups within each degradation type. This plan called for six one-way mixed-effects analyses of variance (ANOVA) with repeated measures on the seven degradation levels within a degradation type (Winer, 1971, pp. 261-268). These analyses made it possible to determine whether mean log geometric mean SQMEs differed as a function of degradation level. If an ANOVA showed a statistically significant trend for the log geometric means to be influenced by changes in degree of degradation, a test for linear trend (Winer, 1971, pp. 296-300) was used. The orthogonal coefficients required for this test were computed according to the method described by Kirk (1968, pp. 513-517) since Winer did not describe a procedure for use with unequal intervals of the independent variable. Tests for higher order trends were used if a linear equation failed to provide a statistically significant fit to the data, or if the linear equation failed to account for more than half the variance. The percentage (%) of variance accounted for by the linear component was estimated roughly by the following formula,

$$\% \text{ Variance} = 100 \left(\frac{SS_{lin}}{SS_{degradation\ level}} \right).$$

The slopes for the log-log functions were examined for statistically significant differences as a function of groups, degradation types, and interactions of groups and degradation types. This plan called for a two-way ANOVA (2 X 3) with repeated measures on degradation types (Winer, 1971, pp. 518-526). In addition to the usual ANOVA summary table, two additional computations were done. The exact probabilities of Type I errors were estimated for all F ratios. A strength of association statistic (eta squared) was computed for significant F ratios (Linton and Gallo, 1975, pp. 335-337). This statistic estimates the proportion of sample variance in the dependent variable (slopes) which can be "explained" by the independent variables (i.e., groups, degradation types, and interactions). Where appropriate, the Newman-Keuls specific comparisons test was used to determine statistically significant differences among specific pairs of means (Linton and Gallo, 1975, pp. 324-327).

Procedures on Visual Training and Screening Data

Computations were done on the visual training and screening data to discover possible differences among groups and subjects with respect to magnitude estimation scaling abilities. This was important since it was desirable to attribute differences in the SQME data to perceptual differences rather than to differences in sophistication with the method of measurement.

Pearson product-moment correlations (Linton and Gallo, 1975, pp. 347-352) were used to check the within-session test-retest reliability of the visual magnitude estimates. Pearson r's were obtained between the magnitude estimates for Trials 2 and 3 for each of the 24 subjects.

To check within-session test-retest reliability of slopes, a two-way ANOVA with repeated measures on Trials 2 and 3 (Winer, 1971, pp. 518-526) was computed for Groups 1 and 2. The analysis made it possible to discover any statistically significant differences as a function of groups, trials, or group-by-trial interactions.

Visual Magnitude Estimation Data

Description

Since subjects were not confined to using a common modulus in making their estimates, typical descriptive statistics were not computed on magnitude estimates or log geometric mean magnitude estimates of circle size. Figure 11, however, shows the mean log geometric means of the modulus-free magnitude estimates plotted as a function of log circle area (in^2) for Group 1, Group 2, and a third group tested during a pilot study of the visual training task (Appendix M). Although it was possible for each subject to assign a different value to each standard stimulus, the mean log geometric means appear to be of similar value at each level of stimulus magnitude. The data points for the three groups also appear to form relatively straight lines with roughly equivalent slopes. The results of the pilot study (Appendix M) showed that a linear equation does, in fact, provide a statistically significant fit of the pilot data.

Table 4 summarizes the mean slopes, standard deviations, and ranges for Groups 1 and 2 and the pilot group. Table 5 summarizes the mean

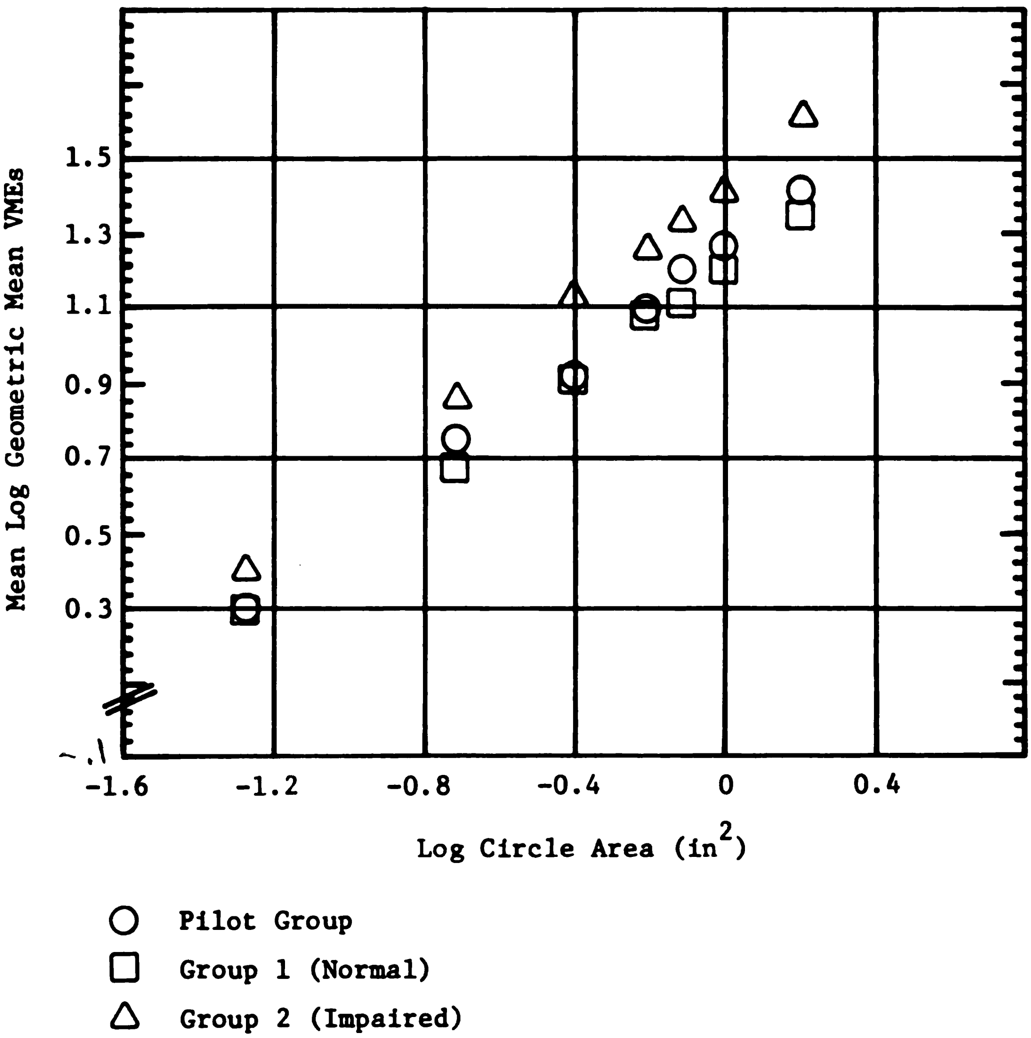

Figure 11. Mean log geometric means of modulus-free visual magnitude estimates (VMEs) plotted as a function of log circle area (in^2) for the pilot group and Groups 1 and 2.

Table 4. Mean slopes, standard deviations, and ranges for the pilot study group, Group 1, and Group 2. Slopes were obtained from least squares solutions for log geometric mean visual magnitude estimates as a function of log circle area (in^2).

	$\overline{X}$	S.D.	Range
Pilot Study Group (n = 12)	.717	0.141	0.459
Group 1 (Normal) (n = 12)	.687	0.118	0.417
Group 2 (Impaired) (n = 12)	.790	0.212	0.667

Table 5. Mean slopes, standard deviations, and ranges for Groups 1 (normal) and 2 (impaired), Trials 2 and 3, and group-by-trial interactions. Slopes were obtained from least squares solutions for log visual magnitude estimates as a function of log circle area (in^2).

	$\overline{X}$	S.D.	Range
Group 1 Trial 2	0.682	0.114	0.417
Group 1 Trial 3	0.691	0.131	0.417
Group 2 Trial 2	0.787	0.227	0.701
Group 2 Trial 3	0.794	0.203	0.632
Trial 2	0.734	0.184	0.705
Trial 3	0.743	0.176	0.683
Group 1	0.686	0.121	0.417
Group 2	0.790	0.211	0.701

slopes, standard deviations, and ranges for Groups 1 and 2, Trials 2 and 3, and the group-by-trial interactions.

Reliability

Table 6 displays the Pearson product-moment correlation coefficients (r) between each subject's Trial 2 and Trial 3 visual magnitude estimates of circle size (in^2). All coefficients were significant beyond the 0.05 level (df = 5; $r_{critical} = 0.754$). Coefficients ranged from 0.98 to 1.00 for the normal hearing subjects (Group 1) and from 0.94 to 1.00 for the hearing-impaired subjects (Group 2). The coefficient of determination for the lowest r in each group was 0.96 for the normal hearing group and 0.88 for the impaired group. Thus, a statistically significant, strong positive relationship existed between each subject's Trial 2 and Trial 3 magnitude estimates. Most of the variance in the magnitude estimates (i.e., over 96% for Group 1 and over 88% for Group 2) could be accounted for by the linear relationship between Trial 2 and Trial 3 estimates. Cumulatively, these results suggest a very high degree of within-subject reliability for both groups.

Table 7 summarizes the results of a two-way analysis of variance in slopes for Groups 1 and 2 with repeated measures on Trials 2 and 3. None of the observed F ratios were significant at the 0.05 confidence level. These results suggest good within-session test-retest reliability of slopes.

Table 8 summarizes the results of an incidental one-way analysis of variance in slopes as a function of groups (i.e., the pilot group and Groups 1 and 2). The observed F ratio was not significant. Thus, there were no systematic differences in the slopes of the three groups on the visual training and screening task.

Table 6. Pearson product-moment correlation coefficients (r) between the visual magnitude estimates of circle size (in^2) for each subject's Trial 2 and Trial 3 stimuli. All coefficients were significant beyond $\alpha \leq 0.05$ (df = 5; $r_{critical} = 0.754$).

Group 1 (Normal)		Group 2 (Impaired)	
Subject	r	Subject	r
1	1.00	13	0.980
2	0.997	14	1.000
3	0.996	15	0.953
4	1.000	16	0.999
5	1.000	17	0.990
6	0.981	18	1.000
7	0.980	19	0.990
8	0.990	20	0.960
9	1.000	21	0.980
10	0.980	22	1.000
11	1.000	23	0.990
12	1.000	24	0.940

Table 7. Analysis of variance in slopes as a function of trials (i.e., Trials 2 and 3) and groups (i.e., Groups 1 and 2). Slopes were obtained from least squares solutions for log visual magnitude estimates as a function of log circle area (in^2).

Source	SS	df	MS	F	F for α=.05	α for $f_{observed}$
Between subjects	1.4277	23	0.0621			
Groups	0.1290	1	0.1290	2.185	4.30*	0.154
Subjects within groups	1.2987	22	0.0590			
Within subjects	0.0570	24	0.0024			
Trials	0.0008	1	0.0008	0.326	4.30*	0.574
Groups X trials	0.0000	1	0.0000	0.004	4.30*	0.948
Trials X subjects within groups	0.0561	22	0.0026			
Total	1.4847	47				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 8. Analysis of variance in slopes as a function of groups (i.e., pilot group and Groups 1 and 2). Slopes were obtained from least squares solutions for log geometric mean visual magnitude estimates as a function of log circle area (in^2).

Source of Variance	SS	df	MS	F	F for α=.05	α for $F_{obs.}$
Between	0.6820536E-01	2	.3410E-01	1.295	3.347*	0.2874
Within	0.8687792	33	.2633E-01			
Total	0.9369845	35				

*From F distribution table (Winer, 1971, pp. 864-869).

Speech Quality Magnitude Estimation Data

Description

Mean log geometric mean modulus-free SQMEs for Groups 1 and 2 are shown in Figure 12 for low-pass filtered stimuli, Figure 13 for high-pass filtered stimuli, and Figure 14 for linearly rectified stimuli. Lines of best fit were plotted by applying the method of least squares. Figure 15 summarizes the lines of best fit for modulus-equalized mean log geometric mean SQMEs for Groups 1 and 2 plotted as a function of log stimulus values for low-pass filtering, high-pass filtering, and linear rectification. Modulus equalization was accomplished graphically by assigning the same arbitrary value to the points on the lines of best fit which represent the mean log geometric mean SQMEs for the standard stimulus. The individually plotted functions for each degradation type appear to be relatively linear in shape, and the magnitudes of the log geometric mean SQMEs are monotonically related to stimulus magnitude (i.e., decreasing stimulus degradation). The composite graph showing lines of best fit, however, suggests that the magnitude of quality perception grows at different rates for the two groups under different degradation types.

Table 9 summarizes the mean slopes, standard deviations, and ranges for listener groups as a function of degradation types. As before, slopes were obtained by applying the method of least squares to the log geometric mean SQMEs and log stimulus values. The smallest mean slope was obtained for Group 2 on low-pass filtered stimuli; the highest mean slope was obtained for Group 1 on linearly rectified stimuli. Figure 16 shows the mean slopes for groups plotted as a function of degradation types. The group with the highest slope varies in a manner which results

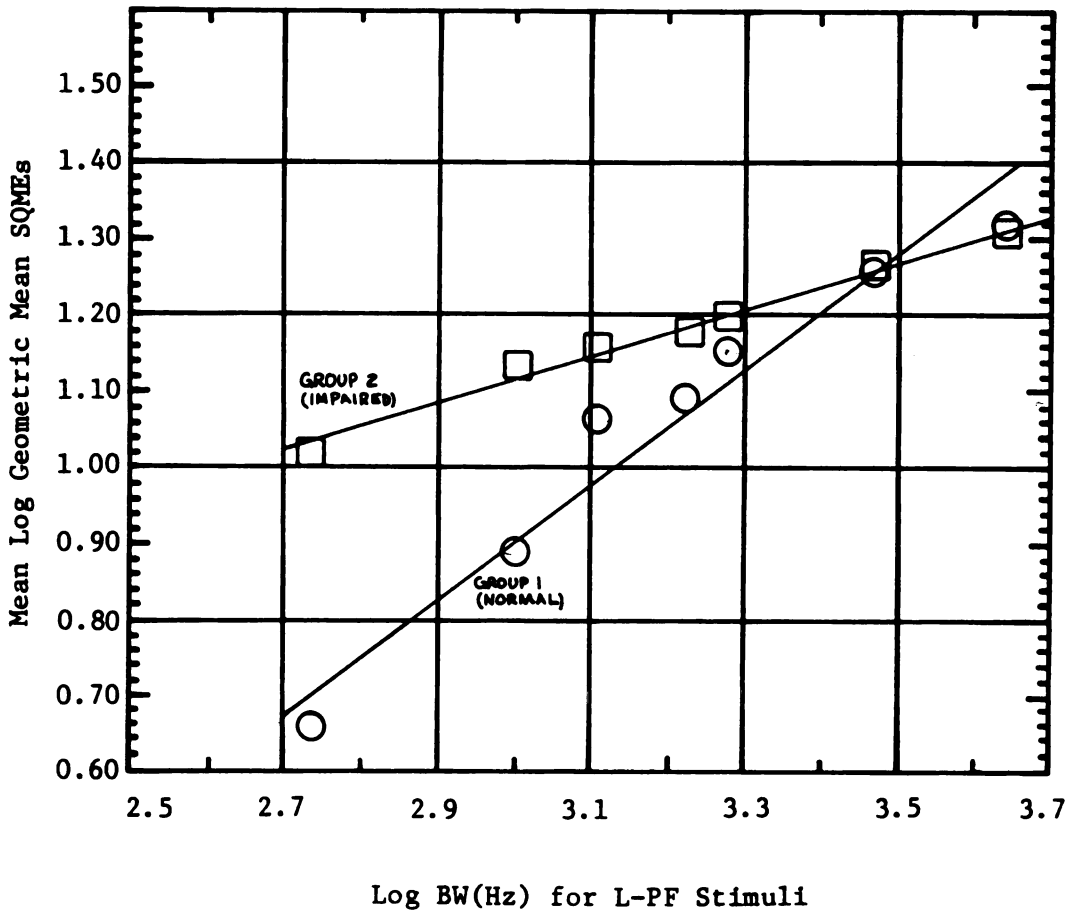

Figure 12. Mean log geometric mean modulus-free SQMEs for Groups 1 and 2 plotted as a function of log frequency bandwidth (Hz) for low-pass filtered stimuli. Lines of best fit were obtained from least square solutions.

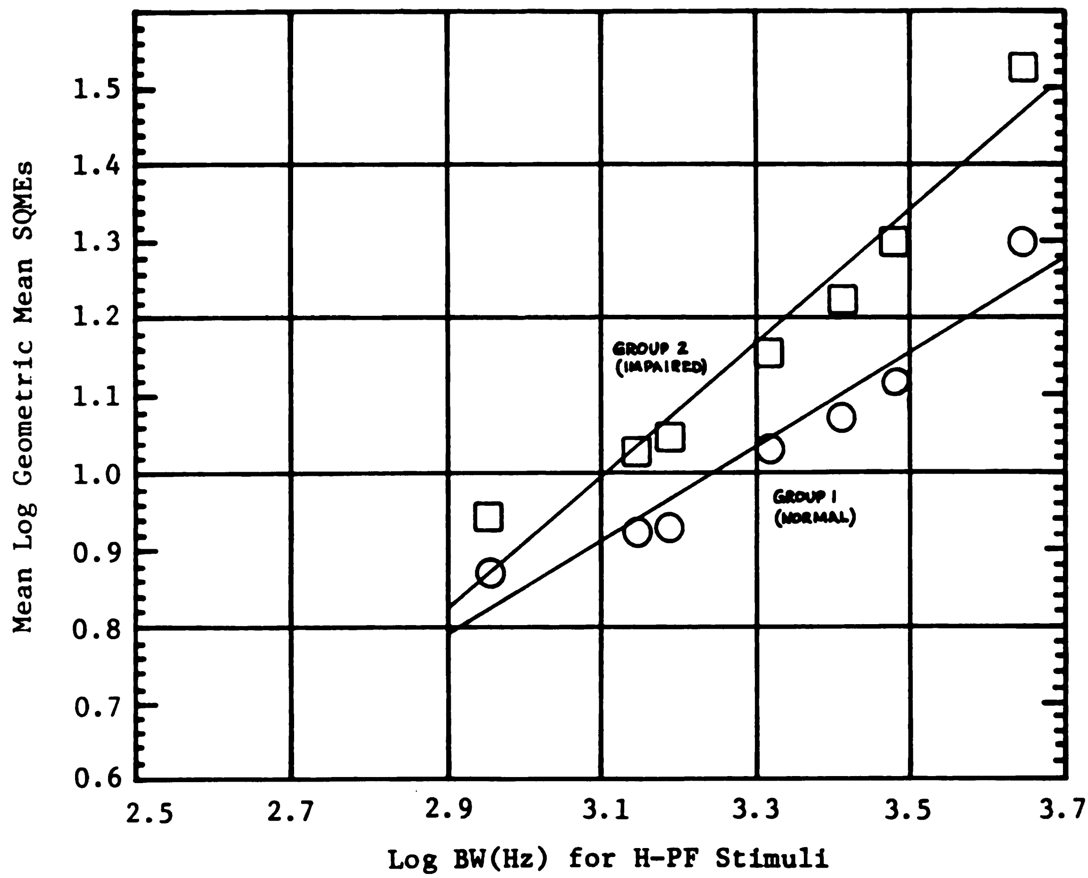

Figure 13. Mean log geometric mean modulus-free SQMEs for Groups 1 and 2 plotted as a function of log frequency bandwidth (Hz) for high-pass filtered stimuli. Lines of best fit were obtained from least squares solutions.

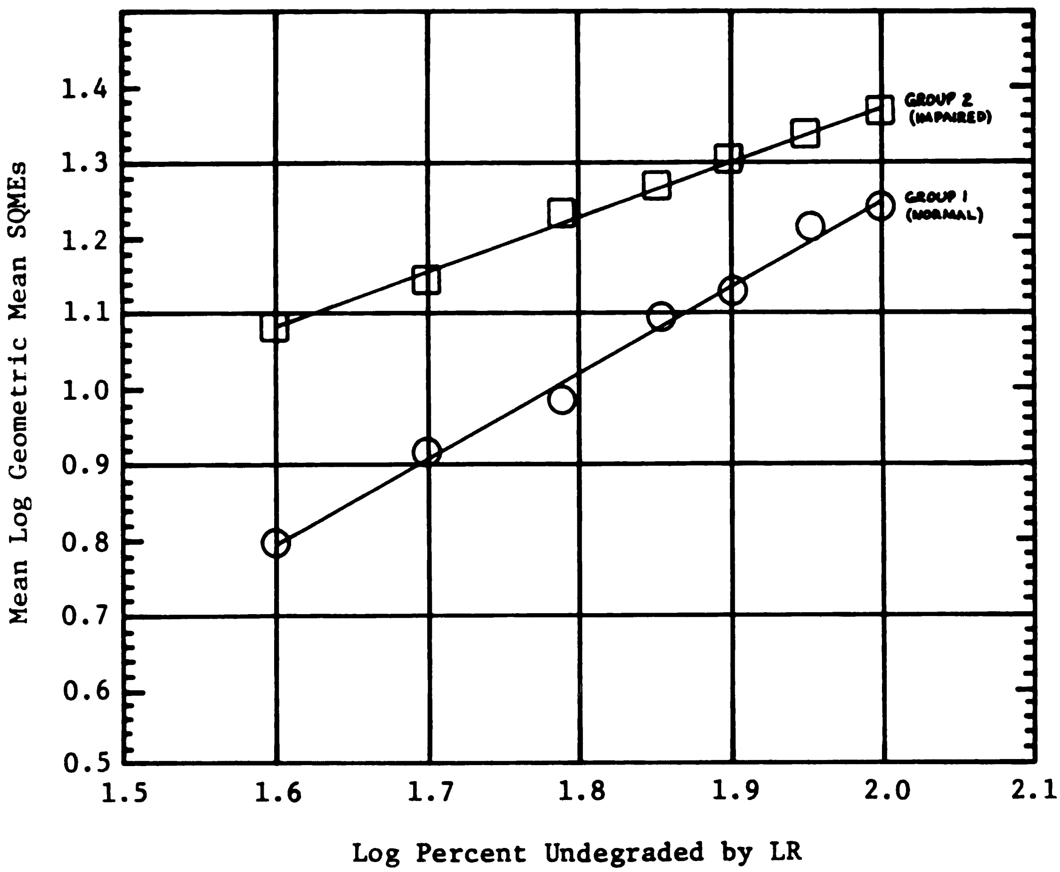

Figure 14. Mean log geometric mean modulus-free SQMEs for Groups 1 and 2 plotted as a function of log percent (%) undegraded by linear rectification. Lines of best fit were obtained from least squares solutions.

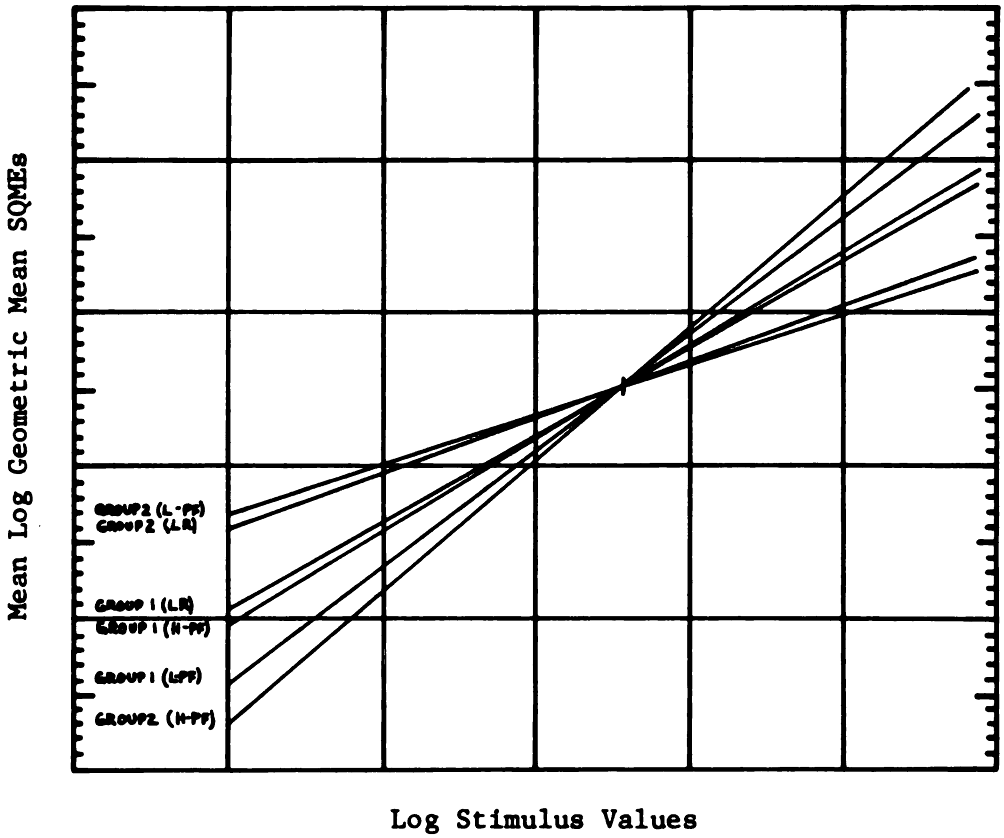

Figure 15. Lines of best fit for modulus-equalized mean log geometric mean SQMEs for Groups 1 (normal) and 2 (impaired) plotted as a function of log stimulus values for low-pass filtering (L-PF), high-pass filtering (H-PF), and linear rectification (LR). Lines of best fit were obtained from least squares solutions. Modulus equalization was accomplished graphically by assigning the same arbitrary value to the points on the lines of best fit which represent the mean log geometric mean SQME for the middle (i.e., the standard) stimulus.

Table 9. Mean slopes, standard deviations, and ranges for Groups 1 (normal) and 2 (impaired) as a function of degradation types. Slopes were obtained from least squares solutions for log geometric mean SQMEs as a function of log stimulus values.

		L-PF			H-PF			LR			All Degradation Types	
Group	$\overline{X}$	S.D.	Range	$\overline{X}$	S.D.	Range	$\overline{X}$	S.D.	Range	$\overline{X}$	S.D.	Range
1	0.744	0.322	0.847	0.620	0.244	0.694	1.144	0.472	1.399	0.836	0.416	1.504
2	0.316	0.286	0.863	0.851	0.256	0.906	0.725	0.507	1.682	0.631	0.425	1.682
1 & 2	0.530	0.369	1.133	0.736	0.271	1.052	0.935	0.525	1.966			

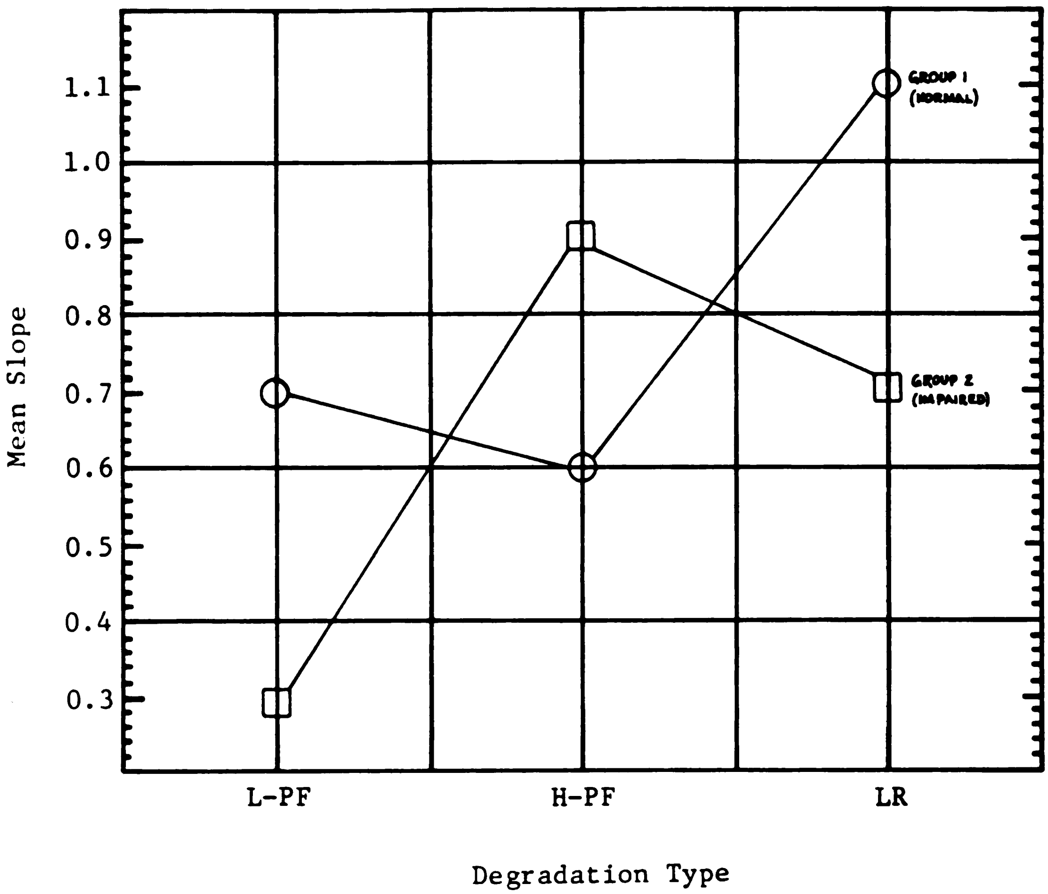

Figure 16. Mean slopes for groups plotted as a function of degradation types. Slopes were obtained from least squares solutions for log geometric mean SQMEs as a function of log degrees of degradation.

in a double transverse interaction. Group 2 shows the largest slope only for high-pass filtered stimuli. In Figure 17 the mean slopes for degradation types are plotted as a function of groups. The slopes for linear rectification are larger than the slopes for low-pass filtering by about the same magnitude for Group 1 and Group 2; the slopes for both degradation types are higher for Group 1 than for Group 2. However, when the slopes for high-pass filtering are considered, a transverse interaction again becomes apparent. For Group 1, the smallest slope among the three degradation types is attributed to high-pass filtering; for Group 2, the greatest slope among the three degradation types is attributed to high-pass filtering.

Reliability

The between-session analyses directly involved the dependent variables, while the within-session analyses involved variables underlying the dependent variables.

Within sessions. Table 10 lists individual subject and group ($n = 12$) "average" correlation coefficients between SQMEs for the first and second trials in Session 1. Group average correlations also are shown graphically in Figure 18. Significant coefficients were those which equaled or exceeded 0.754. For Group 1, significant correlations were found for all but six of the 36 individual subject degradation-type combinations (two under each degradation type). The "average" coefficients for Group 1 ranged from 0.90 for high-pass filtering to 0.93 for low-pass filtering. These averages suggest high within-subject reliability of SQMEs for Group 1 under all degradation types in Session 1. For Group 2, nine of the 36 individual subject correlations were not significant: five under low-pass filtering, one under high-pass

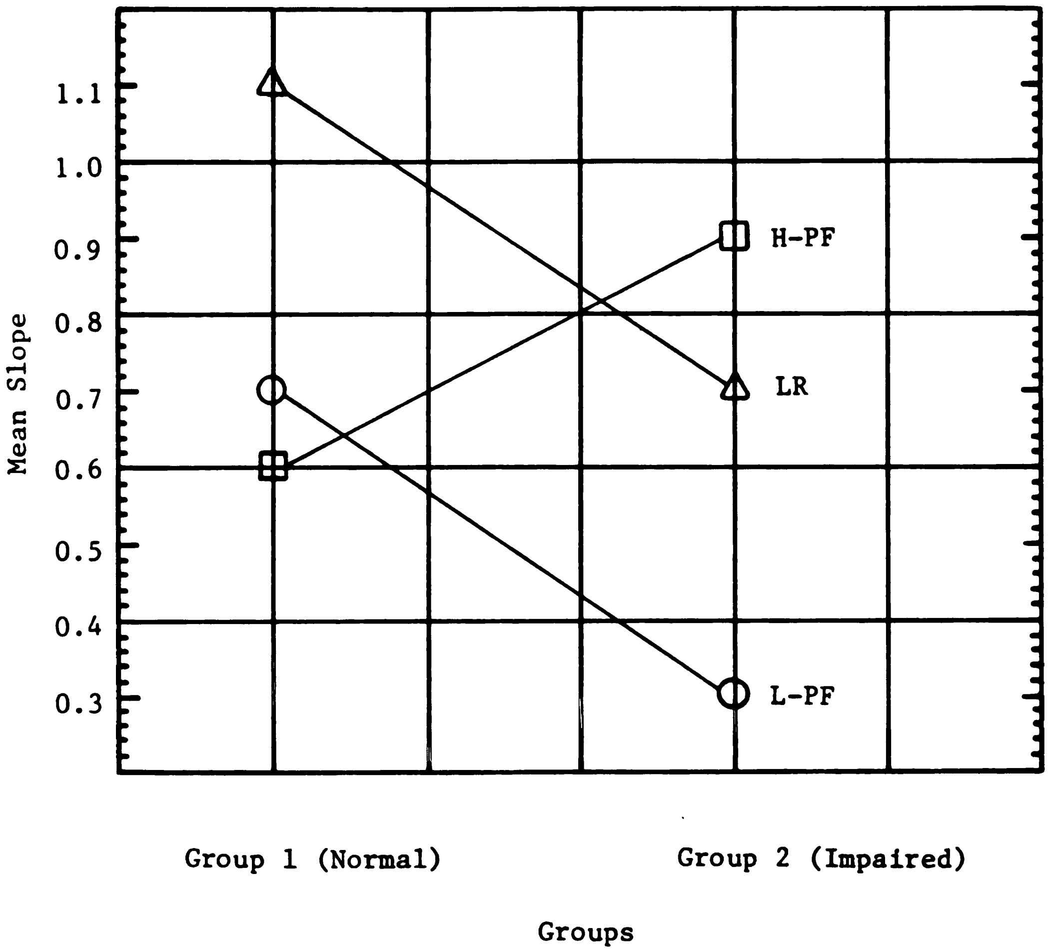

Figure 17. Mean slopes for degradation types plotted as a function of groups. Slopes were obtained from least squares solutions for log geometric mean SQMEs as a function of log degrees of degradation.

Table 10. Within subject correlation coefficients (Pearson r) between seven Trial 1 and Trial 2 SQMEs for each of 24 subjects (i.e., 2 groups of 12) in Session 1. "Average" coefficients were determined for groups within degradation types.

	Degradation Type			
Subject	L-PF	H-PF	LR	$\bar{X}_r^a$
Group 1 (Normal)				
1	0.99	1.00[b]	0.89	
2	0.97	0.65	0.92	
3	0.37	0.92	0.89	
4	0.98	0.97	0.67	
5	0.96	0.77	0.94	
6	0.78	0.83	0.95	
7	0.90	0.90	0.95	
8	0.96	0.85	0.70	
9	0.36	0.86	0.86	
10	0.97	0.95	0.99	
11	0.94	0.88	0.90	
12	0.96	0.52	0.87	
$\bar{X}_r^a$	0.93	0.90	0.91	0.92
Group 2 (Impaired)				
13	0.85	0.99	0.91	
14	0.52	0.98	0.87	
15	-0.54	0.75	0.65	
16	-0.37	0.63	-0.23	
17	0.68	0.93	0.77	
18	0.82	0.98	0.88	
19	0.11	0.97	0.80	
20	0.89	0.89	0.93	
21	0.93	0.86	0.96	
22	0.98	0.997[c]	0.96	
23	0.92	0.94	0.96	
24	0.88	0.91	0.26	
$\bar{X}_r^a$	0.74	0.95	0.84	0.87

[a]"Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963, pp. 680-681).

[b]Changed to r = 0.998 for conservative transformation of r to Z since there is no transformation for r = 1.00.

[c]Interpolated from three decimal place coefficient for conservative transformation of r to Z since there is no transformation for r rounded to 1.00.

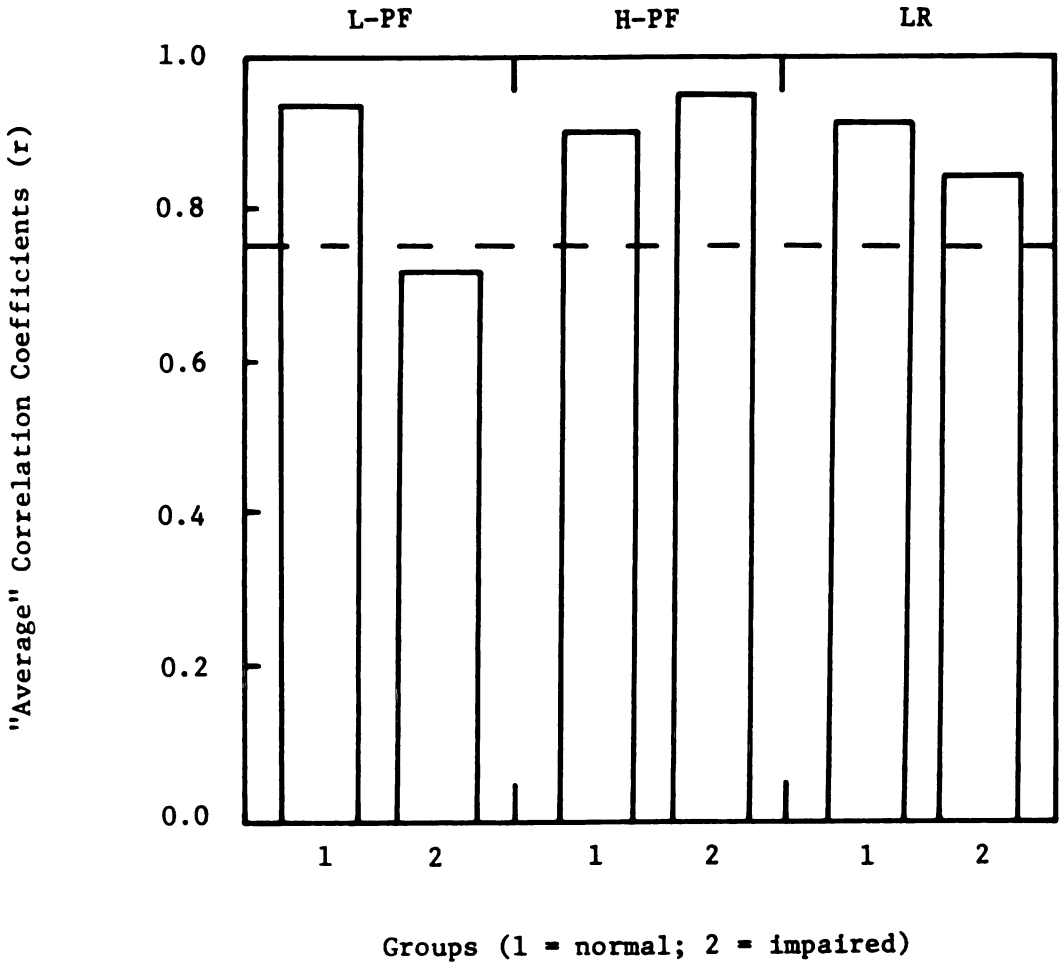

Figure 18. "Average" within-subject test-retest correlation coefficients between seven Trial 1 and Trial 2 SQMEs for two groups of 12 subjects under three degradation types in Session 1. "Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963, pp. 680-681). The dashed horizontal line denotes the significance criterion (R = 0.754).

filtering, and three under linear rectification. "Average" coefficients for Group 2 ranged from 0.74 for low-pass filtering to 0.95 for high-pass filtering. These averages suggest moderate to high within-subject reliability of SQMEs for Group 2 in Session 1. The grand mean coefficients were 0.97 for Group 1 and 0.87 for Group 2. The grand grand mean was 0.89. These results suggest good "average" individual subject reliability of SQMEs for both experimental groups.

Table 11 contains the test-retest correlation coefficients between the Trial 1 and Trial 2 SQMEs for each subject who was tested in two sessions. "Average" coefficients are displayed in Table 11 and in Figure 19 for groups within sessions and degradation types. Significant correlations were those which exceeded 0.754. For Group 1, all individual subject correlations were significant except two for a single subject under low-pass filtering and two for another subject under linear rectification. "Average" coefficients for Group 1 ranged from 0.88 to 0.97, suggesting high within-subject reliability of SQMEs for reliability Group 1 under all degradation types in Sessions 1 and 2. All but four of the Group 2 correlations, three for low-pass filtering and one for linear rectification, were statistically significant. The "average" coefficients for Group 2 ranged from 0.78 to 0.95; these averages suggest that the within-subject, within-session reliability of SQMEs for Group 2 ranged from adequate to high. The grand mean coefficients were 0.93 for Group 1 and 0.90 for Group 2. The grand grand mean across groups was 0.92. These results suggest good average individual subject reliability of SQMEs in the two subgroups.

Between sessions. Table 12 displays the Pearson correlation coefficients between the seven log geometric mean SQMEs for Sessions 1 and 2 for each subject who was tested in two sessions. "Average" coefficients

Table 11. Within subject correlation coefficients (Pearson r) between seven Trial 1 and Trial 2 SQMEs for each of eight subjects (i.e., 2 groups of 4) within Session 1 and Session 2. "Average" coefficients[a] were determined for groups within degradation types.

	Degradation Type						
	L-PF		H-PF		LR		
	Session		Session		Session		
Subject	1	2	1	2	1	2	$\bar{X}_r^a$
Group 1 (Normal)							
3	0.37	0.52	0.92	0.94	0.89	0.87	
4	0.98	0.92	0.97	0.95	0.67	0.62	
10	0.97	0.98	0.95	1.00[b]	0.99	0.92	
11	0.94	0.97	0.88	0.89	0.90	0.95	
$\bar{X}_r$	0.93	0.93	0.94	0.97	0.92	0.88	0.93
Group 2 (Impaired)							
17	0.68	0.51	0.93	0.97	0.77	0.98	
18	0.82	0.65	0.98	0.94	0.88	0.89	
20	0.89	0.91	0.89	0.95	0.93	0.97	
24	0.88	0.86	0.91	0.92	0.26	0.91	
$\bar{X}_r$	0.83	0.78	0.94	0.95	0.79	0.95	0.90

[a]"Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963, pp. 680-681).

[b]Changed to r = 0.998 for conservative transformations of r to Z since there is no transformation for r = 1.

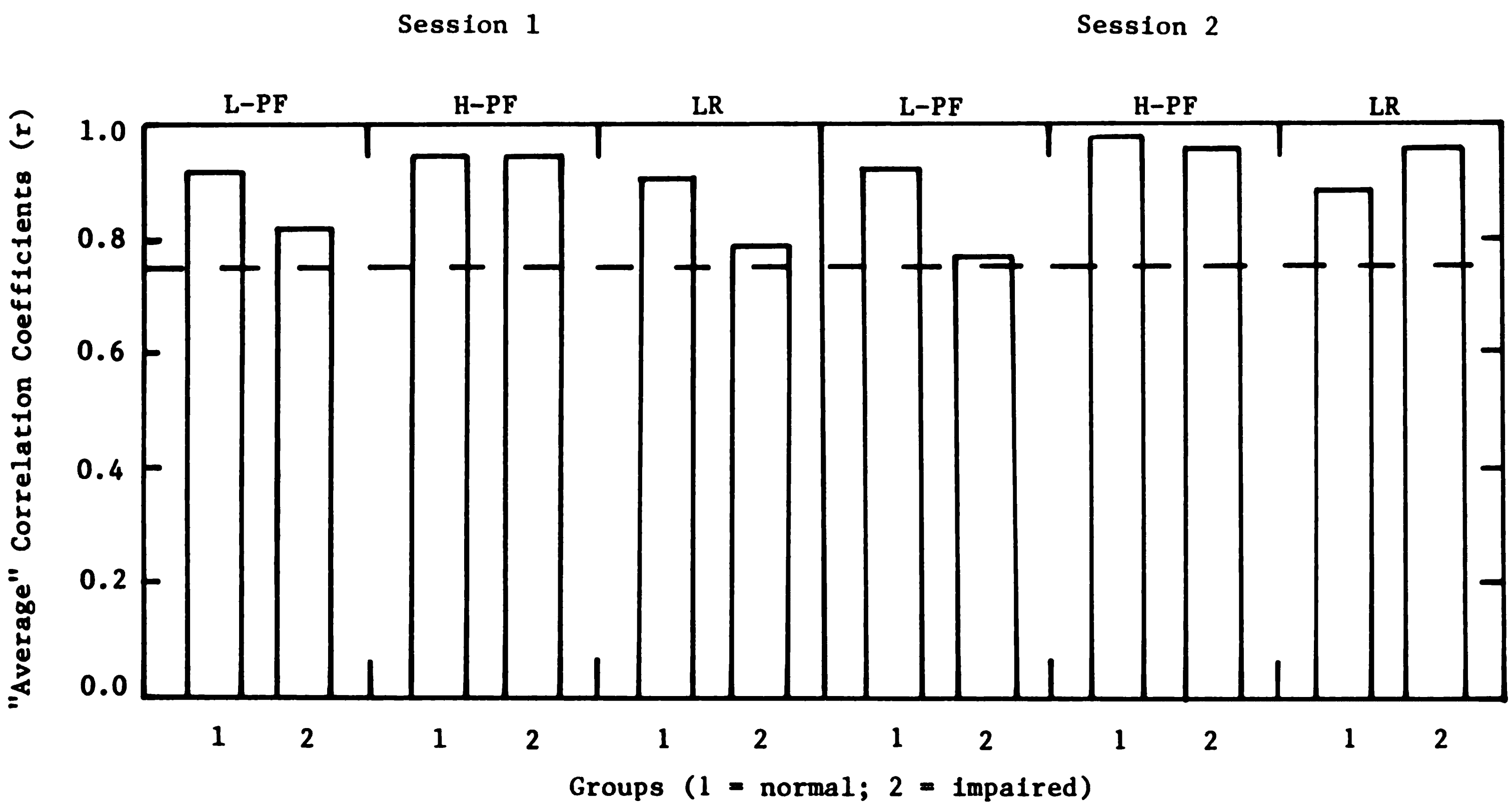

Figure 19. "Average" within-subject test-retest correlation coefficients between seven Trial 1 and Trial 2 SQMEs for two groups of four subjects under three degradation types in Session 1 and Session 2. "Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963, pp.680-681). The dashed horizontal line denotes the significance criterion (r = 0.754).

Table 12. Within subject correlation coefficients (Pearson r) between each of the eight subject's seven log geometric mean SQMEs for Sessions 1 and 2 under the three degradation types. "Average" coefficients were determined for groups within degradation types.

	Degradation Type			
Subject	L-PF	H-PF	LR	$\bar{X}_r^a$
Group 1 (Normal)				
3	0.97	0.88	0.94	
4	0.997[b]	0.93	0.67	
10	0.99	0.96	0.96	
11	0.996[b]	0.96	0.98	
$\bar{X}_r^a$	0.99	0.94	0.94	0.97
Group 2 (Impaired)				
17	0.80	0.97	0.97	
18	0.93	0.96	0.94	
20	0.95	0.99	0.98	
24	0.99	0.91	0.80	
$\bar{X}_r^a$	0.95	0.97	0.95	0.96

[a]"Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963, pp. 680-681).

[b]Derived from three decimal place coefficient for conservative transformation of r to Z since there is no transformation for r rounded to 1.00.

are shown for groups and for groups under degradation types. Significant correlations were those which equaled or exceeded 0.754. All correlations were statistically significant, except one for a single subject in Group 1 under linear rectification. "Average" coefficients ranged from 0.94 to 0.99 with a grand mean of 0.97 for Group 1 and from 0.95 to 0.97 with a grand mean of 0.96 for Group 2. The grand grand mean was 0.96. Thus, the between-session reliability of the log geometric mean SQMEs appears to be very high for both listener groups.

Table 13 and Figure 20 show the test-retest correlation coefficients between Session 1 and Session 2 slopes for the two groups of four subjects under the three degradation types. "Average" coefficients are shown for groups across degradation types, degradation types across groups, and for all data. Statistically significant correlations were those which equaled or exceeded 0.95. However, significance or non-significance may be of questionable value when the number of correlated pairs is only four. Perhaps it is wiser to determine when 50 percent or more of the variance in one set of scores can be accounted for from the other set of scores. This situation exists when r^2, the coefficient of determination (Linton and Gallo, 1975, pp. 329-332), is equal to or greater than 0.50 (i.e., $r_{obs} \geq 0.71$). Under this criterion, meaningful correlations included the Group 1 correlations for low-pass filtering and high-pass filtering and the Group 2 correlations for low-pass filtering and linear rectification. Meaningful "average" correlations consisted of the one for Group 1 across degradation types and the one for low-pass filtered stimuli across groups.

Table 13. Test-retest correlation coefficients (Pearson r) and "average" coefficients between Session 1 and Session 2 slopes for two groups of four subjects under three degradation types. Slopes were obtained from least squares solutions for log geometric mean SQMEs as a function of log stimulus values.

	Degradation Type			
	L-PF	H-PF	LR	$\bar{X}_r^a$
Group 1 (Normal) (n = 4)	0.96	0.73	0.09	0.76
Group 2 (Impaired) (n = 4)	0.88	-0.52	0.91	0.65
$\bar{X}_r^a$	0.93	0.17	0.67	0.71

[a]"Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963, pp. 680-681).

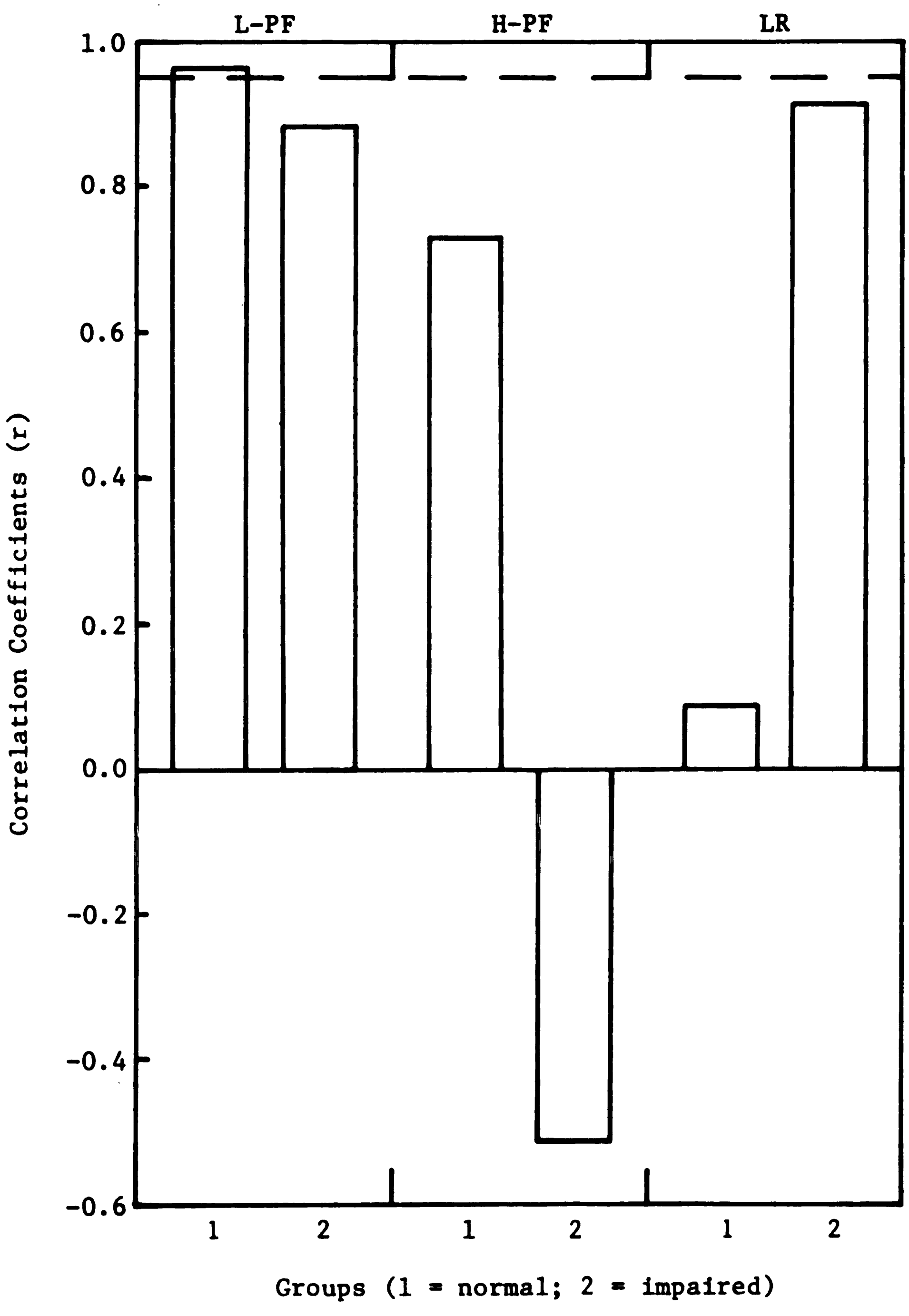

Figure 20. Test-retest correlation coefficients (Pearson r) between Session 1 and Session 2 slopes for two groups of four subjects under three degradation types. The dashed horizontal line denotes the significance criterion (r = 0.95).

Analysis

Results of the primary analyses included those on (1) the presence of trends in log geometric mean SQMEs as a function of log stimulus values, (2) the nature of trends, and (3) the significance of differences in slopes. An alpha level of 0.05 was chosen as the criterion for determining statistical significance.

Presence of trends. Tables 14 through 19 summarize the results of analyses of variance in log geometric mean SQMEs across trials as a function of log degradation levels for each degradation type. Each table shows the results for one of the two experimental groups under one of the three types of degradation. Group 1 data are shown for low-pass filtering (Table 14), high-pass filtering (Table 15), and linear rectification (Table 16); Group 2 data are also shown for low-pass filtering (Table 17), high-pass filtering (Table 18), and linear rectification (Table 19). In each case, the observed F ratio was significant, suggesting a statistically significant trend for the log geometric mean SQMEs to be influenced by changes in degree of degradation.

Nature of trends. Tables 20 through 25 summarize the results of tests to determine whether a significant portion of the trends detected by the analyses of variance could be accounted for by a linear equation. Each table shows the results of a test for linear trend for one of the two experimental groups under one of the three types of degradation. First, Group 1 data are shown for low-pass filtering (Table 20), high-pass filtering (Table 21), and linear rectification (Table 22); then, Group 2 data are shown for low-pass filtering (Table 23), high-pass filtering (Table 24), and linear rectification (Table 25). In each case, the observed F ratio for linear trend was significant, suggesting that a

Table 14. Summary of analysis of variance in log geometric mean SQMEs across trials for Group 1 as a function of log low-pass filtered bandwidth (Hz).

Source	SS	df	MS	F	F for α=.05	α for $F_{obs.}$
Between subjects	10.7219	11	0.9747			
Within subjects	4.5816	72	0.0636			
Degradation levels	3.7916	6	0.6319	52.795	2.242*	$<$ 0.0001
Residual	0.7900	66	0.0120			
Total	15.3035	83				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 15. Summary of analysis of variance in log geometric mean SQMEs across trials for Group 1 as a function of log high-pass filtered bandwidth (Hz).

Source	SS	df	MS	F	F for α=.05	α for $F_{obs.}$
Between subjects	9.9774	11	0.9070			
Within subjects	2.0433	72	0.0284			
Degradation levels	1.5396	6	0.2566	33.618	2.242*	< 0.0001
Residual	0.5038	66	0.0076			
Total	12.0207	83				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 16. Summary of analysis of variance in log geometric mean SQMEs across trials for Group 1 as a function of log percent of undegraded by linear rectification.

Source	SS	df	MS	F	F for α=.05	α for $F_{obs.}$
Between subjects	9.0894	11	0.8263			
Within subjects	2.4580	72	0.0341			
Degradation levels	1.9073	6	0.3179	38.096	2.242*	$<$ 0.0001
Residual	0.5507	66	0.0083			
Total	11.5474	83				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 17. Summary of analysis of variance in log geometric mean SQMEs across trials for Group 2 as a function of log low-pass filtered bandwidth (Hz).

Source	SS	df	MS	F	F for α=.05	α for $F_{obs.}$
Between subjects	8.2609	11	0.7510			
Within subjects	1.2846	72	0.0178			
Degradation levels	0.6638	6	0.1106	11.763	2.242*	$<$ 0.0001
Residual	0.6208	66	0.0094			
Total	9.5455	83				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 18. Summary of analysis of variance in log geometric mean SQMEs across trials for Group 2 as a function of log high-pass filtered bandwidth (Hz).

Source	SS	df	MS	F	F for α=.05	α for $F_{obs.}$
Between subjects	8.2438	11	0.7494			
Within subjects	3.5472	72	0.0493			
Degradation levels	2.8280	6	0.4713	43.250	2.242*	$<$ 0.0001
Residual	0.7193	66	0.0109			
Total	11.7910	83				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 19. Summary of analysis of variance in log geometric mean SQMEs across trials for Group 2 as a function of log percent undegraded by linear rectification.

Source	SS	df	MS	F	F for α=.05	α for $F_{obs.}$
Between subjects	7.8678	11	0.7153			
Within subjects	1.3085	72	0.0182			
Degradation levels	0.7709	6	0.1285	15.773	2.242*	$<$ 0.0001
Residual	0.5376	66	0.0081			
Total	9.1764	83				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 20. Results of test for linear trend in log geometric mean SQMEs for Group 1 as a function of log bandwidths (Hz) of low-pass filtered stimuli.

Source of Variance	SS	df	MS	F	F for α=0.05	α for $F_{obs.}$
Linear trend	3.62	1	3.62	362	4.0147	< 0.0001
Deviation from linear trend	0.96	71	0.01			

*Interpolated from F distribution table (Winer, 1971, pp. 864-869).

Table 21. Results of test for linear trend in log geometric mean SQMEs for Group 1 as a function of log bandwidths (Hz) of high-pass filtered stimuli.

Source of Variance	SS	df	MS	F	F for $\alpha=0.05$	α for $F_{obs.}$
Linear trend	1.43	1	1.43	143	4.0147	< 0.0001
Deviation from linear trend	0.61	71	0.01			

*Interpolated from F distribution table (Winer, 1971, pp. 864-869).

Table 22. Results of test for linear trend in log geometric mean SQMEs for Group 1 as a function of log-percent (%) undegraded by linear rectification.

Source of Variance	SS	df	MS	F	F for α=0.05	α for $F_{obs.}$
Linear trend	1.89	1	1.89	189	4.0147*	< 0.0001
Deviation from linear trend	0.57	71	0.01			

*Interpolated from F distribution table (Winer, 1971, pp. 864-869).

Table 23. Results of test for linear trend in log geometric mean SQMEs for Group 2 as a function of log bandwidths (Hz) of low-pass filtered stimuli.

Source of Variance	SS	df	MS	F	F for α=0.05	α for $F_{obs.}$
Linear trend	0.65	1	0.65	65	4.0147*	< 0.0001
Deviation from linear trend	0.63	71	0.01			

*Interpolated from F distribution table (Winer, 1971, pp. 864-869).

Table 24. Results of test for linear trend in log geometric mean SQMEs for Group 2 as a function of log bandwidth (Hz) of high-pass filtered stimuli.

Source of Variance	SS	df	MS	F	F for α=0.05	α for $F_{obs.}$
Linear trend	2.69	1	2.69	269	4.0147*	< 0.0001
Deviation from linear trend	0.86	71	0.01			

*Interpolated from F distribution table (Winer, 1971, pp. 864-869).

Table 25. Results of test for linear trend in log geometric mean SQMEs for Group 2 as a function of log percent (%) undegraded by linear rectification.

Source of Variance	SS	df	MS	F	F for α=0.05	α for $F_{obs.}$
Linear trend	0.76	1	0.76	76	4.0147*	< 0.0001
Deviation from linear trend	0.55	71	0.01			

*Interpolated from F distribution table (Winer, 1971, pp. 864-869).

linear equation provides a statistically significant fit to the data. Also, the linear component of the trend in each case accounted for a very high percentage (93-99%) of the variance in log geometric mean SQMEs due to log degradation levels. Thus, tests for higher order trends were not employed. Table 26 summarizes the approximate percentages of variance that can be attributed to the linear component of each trend.

Differences in slopes. Table 27 shows the results of a two-way analysis of variance in slopes as a function of the two listener groups (Group 1 and Group 2) and three degradation types (low-pass filtering, high-pass filtering, and linear rectification). Measures were repeated on degradation types. Significant F ratios are those which exceed the critical values of F at an alpha level of 0.05. The main effect of groups was not significant. The main effect for degradation types, however, was significant, suggesting a statistically significant difference in slopes as a function of degradation types. The group-by-degradation type interaction was also significant, indicating that the effects of degradation types and groups were interdependent. Significant effects in the analysis of variance were followed by a Newman-Keuls specific comparison test to evaluate differences within specific pairs of means.

Table 28 shows the results of a Newman-Keuls specific comparison test on pairs of mean slopes for the three degradation types. All differences among the mean slopes for low-pass filtering, high-pass filtering, and linear rectification were statistically significant.

Table 29 shows the results of a Newman-Keuls specific comparison test on pairs of mean slopes for the group-by-degradation type interaction. Nine comparisons revealed differences that were significant beyond the 0.05 level. The pattern of the important differences may be described as follows. The mean slopes for Group 1 differed from those for Group 2

Table 26. Approximate percentages of variance in log geometric mean SQMEs due to log degradation levels that could be accounted for by a linear equation. Percentages (%) are shown as a function of group and degradation type.

Group	Degradation Type	% Variance
1 (Normal)	Low-pass filtering	96
1	High-pass filtering	93
1	Linear rectification	99
2 (Impaired)	Low-pass filtering	98
2	High-pass filtering	95
2	Linear rectification	99

Table 27. Results of a two-way analysis of variance in slopes as a function of two listener groups (Group 1 and Group 2) and three degradation types (low-pass filtering, high-pass filtering, and linear rectification). The slopes were obtained by applying the method of least squares to the log geometric mean SQMEs and log degrees of degradation.

Source	SS	df	MS	F	F for α=0.05	α for $F_{obs.}$
Between subjects	7.1099	23	0.3091			
Groups	0.7612	1	0.7612	2.638	4.300*	0.119
Subjects within groups	6.3487	22	0.2886			
Within subjects	6.0174	48	0.1254			
Degradation types	1.9660	2	0.9830	18.496	3.214*	$<$0.0001
Groups X degradation types	1.7129	2	0.8565	16.115	3.214*	$<$0.0001
Degradation types X subjects within groups	2.3385	44	0.0531			
Total	13.1273	71				

*From F distribution table (Winer, 1971, pp. 864-869).

Table 28. Results of the Newman-Keuls specific comparison test on pairs of mean slopes for the three degradation types. Critical values are given for all possible ranges of means spanned (i.e., two or three means). A difference between any two means is significant when it exceeds the appropriate critical value (CV) for $\alpha = 0.05$. The number of means spanned is equal to k.

MEANS	L-PF 0.530	H-PF 0.736	LR 0.935	CV_k
L-PF		0.206*	0.405*	$CV_2 = -.1345$
H-PF			0.199*	$CV_3 = 0.1618$

*Denotes a significant difference between a pair of means.

Table 29. Results of the Newman-Keuls specific comparison test on pairs of mean slopes for the group-by-degradation type interaction (2 groups X 3 degradation types = 6 means). Critical values are given for all possible ranges spanning from two to six means. A difference between any two means is significant when it exceeds the appropriate critical value (CV) for $\alpha = 0.05$. The number of means is equal to k.

MEANS	Group 2 X L-PF 0.316	Group 1 X H-PF 0.620	Group 2 X LR 0.725	Group 1 X L-PF 0.744	Group 2 X H-PF 0.851	Group 1 X LR 1.144	CV_k
Group 2 X L-PF		0.304*	0.409*	0.428*	0.535*	0.828*	$CV_2 = 0.190$
Group 1 X H-PF			0.105	0.124	0.231	0.524*	$CV_3 = 0.228$
Group 2 X LR				0.019	0.126	0.419*	$CV_4 = 0.251$
Group 1 X L-PF					0.107	0.400*	$CV_5 = 0.268$
Group 2 X H-PF						0.293*	$CV_6 = 0.280$

*Denotes a significant difference between a pair of means.

as a function of low-pass filtering and linear rectification, but not as a function of high-pass filtering. Within Group 1, the mean slope for low-pass filtering differed from the mean for linear rectification, but not the mean for high-pass filtering; the mean slopes for high-pass filtering and linear rectification, however, also differed. Within Group 2, the mean slopes for low-pass filtering differed from the means for high-pass filtering and linear rectification; the mean slopes for high-pass filtering and linear rectification, however, did not differ. Mean slopes for groups (Figure 16, p. 67) varied as a function of degradation types in a manner which resulted in a double transverse interaction. The mean Group 1 slope was higher than the mean Group 2 slope for low-pass filtering, lower than the mean Group 2 slope for high-pass filtering, and higher than the mean Group 2 slope for linear rectification. This pattern may partially account for the failure to find a significant main effect for groups.

Effects found to be significant by the analysis of variance were also followed by a strength of association statistic. Eta squared (η^2) was computed to estimate the proportion of variance in slopes that could be accounted for by main effects and the group-by-degradation type interaction effect. For degradation types, η^2 was 0.1498, suggesting that approximately 15% of the variance in slopes could be attributed to degradation type. For the group-by-degradation type interaction, η^2 was 0.1305, suggesting that approximately 13% of the variance in slopes could be attributed to the interaction. The relatively low percentages of variance in slopes accounted for the significant effects suggests a relatively mild experimental effect in both cases.

Summary

Visual Magnitude Estimation Data

The log-log plots for modulus-free magnitude estimates as a function of circle size were relatively linear and had roughly equivalent slopes. High within-session test-retest correlation coefficients were obtained for the magnitude estimates. The nonsignificant results of a two-way analysis of variance in slopes as a function of groups and trials suggested good within-session reliability of slopes. The results of a one-way analysis of variance showed no statistically significant differences in slopes for the normal hearing experimental group (Group 1), the sensorineural hearing loss group (Group 2), and a separate pilot group of normal hearing subjects.

Speech Quality Magnitude Estimation Data

Reliability. The dependent variables considered in the experimental questions were log geometric means and slopes. Since log geometric means were obtained for SQMEs across trials, their test-retest reliability was dependent upon the raw data. Since individual slopes were obtained for the lines which relate log geometric mean SQMEs to log degradation levels, their reliability was dependent upon log geometric mean SQMEs. It should be noted that while individual subject correlations were obtained for SQMEs and log geometric mean SQMEs, within-subject correlations were obtained for slopes.

"Average" within-session reliability of individual subject SQMEs for groups (n = 12) within degradation types was generally high for Group 1 and ranged from moderate to high for Group 2. This same pattern prevailed for the subgroups (n = 4) in Session 1 and Session 2. Overall, the "average" individual subject reliability of SQMEs was high for both data sets (n = 12, n = 4).

Between-session reliability of log geometric mean SQMEs for individual subjects was very high for both subgroups (n = 4) within each degradation type.

The between-session reliability of slopes showed considerable variability as a function of group-by-degradation type interaction. The most reliable slopes appeared to be the Group 1 slopes for low-pass and high-pass filtering and the Group 2 slopes for low-pass filtering and linear rectification.

Trend analyses. The results of six one-way analyses of variance showed statistically significant differences in the log geometric mean SQMEs for each of the two groups as a function of degradation levels for each of the three degradation types. After each analysis of variance, results of tests for linear trend showed that a linear equation provides a statistically significant fit to the data and accounts for a large portion of the variance due to log degradation levels. Thus, a statistically significant linear trend exists for each of the two groups under each of three degradation types.

Slope analyses. The results of a two-way analysis of variance in slopes as a function of listener groups and degradation types showed statistically significant differences for degradation types and group-by-degradation type interaction, but not for groups. Differences among means for degradation types were significant for all comparisons. Mean slopes for groups varied as a function of degradation types such that a double transverse interaction occurred. The mean slope for Group 1 was greater than the mean slope for Group 2 for low-pass filtering and linear rectification but not for high-pass filtering. Mean slopes for degradation types varied as a function of groups such that a transverse interaction resulted. Mean slopes for low-pass filtering and linear rectification were higher for Group 1 than for Group 2. The mean slope for low-pass filtering was less than the mean slope for linear rectification by approximately the same amount for Group 1 and Group 2; so, no interaction was evident when only these two types of degradation were considered. When all three types of degradation were considered, however, high-pass filtering yielded the lowest mean for Group 1 and the highest mean for Group 2, resulting in a transverse interaction. There were no statistically significant differences among the mean slopes for Group 1 - low-pass filtering, Group 1 - high-pass filtering, Group 2 - high-pass filtering, and Group 2 - linear rectification.

CHAPTER IV

DISCUSSION

Introduction

Chapter III described the results produced by 12 normal hearing subjects (Group 1) and 12 sensorineurally impaired hearing subjects (Group 2) on a visual training and screening task and in a speech quality magnitude estimation (SQME) experiment. On the visual task modulus-free magnitude estimates were obtained as a function of circle and square size. In the SQME experiment, modulus-free SQMEs were obtained as a function of seven degrees of three types of signal degradation (low-pass filtering, high-pass filtering, and linear rectification).

Two dependent variables were derived from the visual magnitude estimates: (1) log visual magnitude estimates and (2) the slopes of the lines relating log visual magnitude estimates to log circle size. The visual magnitude estimates were analyzed to determine the strength of the relationship between the Trial 2 and Trial 3 estimates. The mean log visual magnitude estimates for Group 1 and Group 2 were plotted as a function of log circle size for comparison with similar data from a pilot study (Appendix M). The slopes of the log-log functions for Groups 1 and 2 were analyzed for significant differences as a function of groups, trials, and group-by-trial interaction. The slopes of the log-log functions for Groups 1 and 2 and the pilot group were analyzed for statistical significance as a function of groups.

Two dependent variables were also derived from the individual SQMEs: (1) log geometric mean SQMEs and (2) slopes for the lines relating log geometric mean SQMEs to log stimulus values (degrees of degradation).

The log geometric mean SQMEs were examined for statistical significance as a function of degrees of degradation for each listener group and degradation type. When differences were found, these data were further analyzed to identify the lowest order equation required to provide a statistically significant fit to the data. The slopes were analyzed to find statistical significance as a function of listener group, degradation type, and group-by-degradation type interaction.

The findings and their implications are discussed below.

Visual Magnitude Estimation Task

Findings

A statistically significant, strong positive correlation ($r = 0.94$) existed between each subject's Trial 2 and Trial 3 magnitude estimates during the screening task. The slopes relating the log magnitude estimates to log circle size did not differ as a function of groups, trials, or group-by-trial interaction, suggesting good within-session reliability of slopes. The slopes relating log geometric mean magnitude estimates across trials to log circle size did not differ for Group 1, Group 2, or the normal hearing pilot group. The log-log functions for these three groups were similar and were linear in nature. The mean slopes for the three groups were in good agreement with the 0.7 slope reported for similar data by S. S. Stevens (1975). Overall, the reliability of the visual data may be characterized as excellent.

Implications for the SQME Experiment

The strong individual subject correlations suggest that each subject was able to reliably estimate visual magnitude during the visual training and screening tasks. The failure to find within-session differences in slopes as a function of groups, trials, and group-by-trial interaction suggests that: (1) slopes are also replicable within a session and (2) there were no systematic differences in the performance of Group 1 and Group 2 on the visual magnitude estimation task.

The failure to find differences in slopes for Group 1, Group 2, and a normal hearing pilot group suggests that slopes are replicable across time as well as groups. This implication is supported by the agreement of the group slopes with results reported by S. S. Stevens (1975). The fact that the log-log functions are linear suggests that the data represent a power function (S. S. Stevens, 1957).

Collectively, the visual magnitude estimation data suggest that any systematic differences in performance on magnitude estimation tasks which immediately follow the visual tasks are probably due to factors other than lack of skill in magnitude estimation.

SQME Experiment

Reliability

Within sessions. Pearson product-moment correlation coefficients (r) were used to assess the within-session reliability of the raw data (SQMEs). Pearson r's were determined for each of the 12 subjects in each experimental group under each degradation type in Session 1 and Session 2. For the normal hearing listeners "average" individual subject correlations were consistently high ($r = 0.90$ to $r = 0.93$) for the experimental group ($n = 12$) and ranged from high ($r = 0.88$) to extremely high ($r = 0.97$) for the subgroup ($n = 4$). For the hearing impaired listeners "average" individual subject correlations ranged from moderate to high for the experimental group ($r = 0.74$ to $r = 0.95$) and for the subgroup ($r = 0.78$ to $r = 0.95$).

The "average" within-session reliability of SQMEs for individual subjects appears to be good. This reliability is important primarily because the dependent variables (i.e., the log geometric mean SQMEs and slopes) were derived from the SQMEs. The between-session agreement of the log geometric means reflects the reliability of the SQMEs.

Between sessions. The between-session agreement of log geometric mean SQMEs was checked for each of the four subjects in the two subgroups that returned to participate in Session 2. "Average" individual subject correlations between log geometric mean SQMEs for Session 1 and Session 2 were positive for all degradation types; the "average" coefficients ranged from high to extremely high for Group 1 ($r = 0.94$ to $r = 0.99$) and for Group 2 ($r = 0.95$ to $r = 0.97$).

Assuming that the individual ratio scales for the SQMEs did not differ drastically from Session 1 to Session 2, it seems reasonable that the between-session correlations of the log geometric mean SQMEs are stronger than the correlations of SQMEs. This is to be expected since the within-session correlations of SQMEs were strong and since the log geometric mean reduces the effects of extreme scores. The between-session correlations of log geometric mean SQMEs appear to reflect excellent "average" individual subject agreement over time.

The between-session relationship of Session 1 and Session 2 slopes was checked for each subgroup ($n = 4$) under each of the three types of degradation. For Group 1, the strength of the relationship was essentially nil ($r = 0.09$) for linearly rectified stimuli, moderate ($r = 0.73$) for high-pass filtered stimuli, and very high ($r = 0.96$) for low-pass filtered stimuli. For Group 2, the relationship was a weak to moderate one ($r = -0.52$) for high-pass filtered stimuli and a strong positive one for linearly rectified stimuli ($r = 0.91$) and low-pass filtered stimuli (r = 0.88). "Average" correlations across degradation types were moderate for Group 1 (r = 0.76) and Group 2 (r = 0.65). "Average" correlations across groups were nil (r = 0.17) for high-pass filtered stimuli, moderate (r = 0.67) for linearly rectified stimuli, and strong (r = 0.93) for low-pass filtered stimuli.

Why did between-session correlations show poor test-retest agreement for Group 1 slopes for linearly rectified stimuli and of Group 2 slopes for high-pass filtered stimuli? All subjects were trained in magnitude estimation and demonstrated skill in using this procedure with visual stimuli and the remaining auditory stimuli. The order of stimulus presentations was counterbalanced to reduce order effects. Thus, it seems reasonable to speculate that the poor test-retest agreement is due to perceptual difficulty during one or both listening sessions.

It is likely that Group 2 listeners had perceptual difficulty (reflected in poor reliability) with high-pass filtered stimuli because they had high frequency hearing losses and were unable to benefit from much of the high frequency portions of the signals presented. On the other hand, it is possible that the normal hearing listeners had difficulty with linearly rectified stimuli because the amount of information processed and the signal intelligibility were, for them, relatively unaffected by changes in percentage of total harmonic distortion (THD).

Log Geometric Mean SQMEs

The first experimental question asked whether there was a statistically significant trend for the log geometric mean SQMEs to be influenced by changes in degrees of degradation for each listener group under each degradation type. On the basis of the results obtained, it can be said that such a trend does exist. Statistically significant differences were found in the log geometric means for Group 1 and Group 2 as a function of degree of degradation for each of the three degradation types.

The second experimental question was concerned with determining the lowest order equation required to provide a satisfactory fit to the data for each listener group under each degradation type. Examination of test results for linear trend suggested that a first order equation provides a statistically significant fit to and accounts for a large portion of the variance in log geometric means as a function of log degradation levels. Thus, it can be said that a linear trend exists for Group 1 and Group 2 under each type of degradation.

In general, the log geometric mean SQMEs are positively correlated with changes in bandwidth due to low-pass and high-pass filter cutoffs and with percent undegraded by linear rectification (% undegraded = 100% - % THD). As a group phenomenon, at least, log geometric mean SQMEs appear to be linearly sensitive to changes in the log electroacoustic characteristics of the signal. These results tend to support those of paired comparison studies (Jeffers, 1960; Zerlin, 1962; Witter and Goldstein, 1971; Smaldino, 1974; and Chial and Daniel, 1977) which found that hearing aids with better electroacoustic characteristics are preferred over those with poorer electroacoustic characteristics.

The linear nature of the log-log functions indicates that the data represent a power function (S. S. Stevens, 1957). S. S. Stevens (1975) suggested that the slopes of such functions represent a dependent variable which may indicate different perceptual events for different stimuli. According to the power law (S. S. Stevens, 1957), SQMEs should increase as the power or exponent of the signal magnitude (i.e., degree of degradation) increases. In other words, $\psi = \kappa\phi^{\beta}$ where ψ = subjective magnitude (SQME), κ = a scaling factor equal to the intercept, ϕ = stimulus magnitude (i.e., degree of degradation), and β represents the slope of a log-log function.

Do the observed power functions represent prothetic or metathetic continua of perceptual magnitude? Prothetic continua are generally concerned with decisions of quantity, degree, or how much, whereas metathetic continua are generally concerned with decisions of quality, place, what, or where (S. S. Stevens, 1957, 1975). Although power-law behavior is generally exhibited by prothetic continua and not by metathetic continua, the existence of power functions is not independently sufficient to distinguish between the two types of continua (Marks, 1974; S. S. Stevens, 1975; and Gescheider, 1976). In order to draw a distinction between prothetic and metathetic continua, S. S. Stevens (1957) suggested four functional criteria concerned with (1) subjective size of the just noticeable difference (JND), (2) the form of category rating-scales, (3) the time-order error, and (4) hysteresis effects due to stimulus presentation order. In the current study, neither JNDs nor category judgments were obtained, and hysteresis effects were minimized by randomizing degrees of degradation. Thus, the power functions obtained in the present study cannot be readily characterized in terms of these criteria. According to S. S. Stevens (1957, 1975), a time-order error exists on prothetic continua such that a comparison stimulus (second) is judged to be greater than the standard stimulus (first) when the two are equal. The time-order error does not exist on metathetic continua. Table 30 shows the percentages of times the speech quality magnitude of the comparison stimulus equal to the standard stimulus was judged to be less than, equal to, and greater than that of the standard stimulus for each group under each degradation type. With the exception of Group 2 judgments under linear rectification, the quality of the comparison stimulus was judged to be

Table 30. Percentages of times the speech quality magnitude of the comparison stimulus equal to the standard stimulus was judged less than, equal to, or greater than that of the standard stimulus. Percentages are based on the magnitudes of the 24 SQMEs (2 passages judged by 12 subjects) produced by Groups 1 and 2 for the standard degradation level under each degradation type.

	Perceived Quality of the Comparison Stimulus Relative to that of an Equal Standard Stimulus	Degradation Types		
		L-PF	H-PF	LR
	Less	37.5%	37.5%	12.5%
Group 1	Equal	45.8%	41.7%	62.5%
	Greater	16.7%	20.8%	25.05%
	Less	16.7%	33.3%	16.7%
Group 2	Equal	45.8%	41.7%	41.7%
	Greater	37.5%	25.0%	41.7%

equal to that of the standard stimulus more often than it was judged to be less or greater. For Group 2 judgments under linear rectification, the quality of the comparison stimulus equal to that of the standard was judged to be equal to and greater than the quality of the standard an equal number of times. Only in the case of Group 1 under linear rectification, however, did the judged-equal category represent 50 percent or more of the judgments. Thus, the time-order error appears to be of little help in classifying the continua for SQMEs.

Slopes

Mathematically, the slope can be defined as change in the variable on the Y axis divided by the change in the variable on the X axis. When log geometric mean SQMEs are plotted as a function of log degree of degradation, the slope represents increase in perceptual values relative to increase in degree of degradation. If the value of the slope is one, change in log geometric mean SQMEs and change in log degree of degradation are equal and the function creates a 45° angle with the X axis. If the value of the slope is greater than one, change in log geometric mean SQMEs is greater than change in log degree of degradation, and the function creates an angle greater than 45° with the X axis. If the value of the slope is less than one, change in log geometric mean SQMEs is less than change in log degree of degradation, and the function creates an angle of less than 45° with the X axis.

The third experimental question related to the existence of statistically significant differences among the slopes of the log-log functions as a function of listener group, degradation type, and the interaction of listener group and degradation type. The results of the study showed differences in slopes as a function of degradation types and group-by-degradation type interaction, but not as a function of groups. The means, standard deviations, and ranges of the slopes are shown in Table 9 (p. 66) as a function of groups and degradation types.

Effects of degradation type for Group 1 (normal). Consider the mean slopes for Group 1 separately for each degradation type. The change in log geometric mean SQMEs relative to the change in log degree of degradation was essentially the same for low-pass and high-pass filtered stimuli, but was significantly greater for linearly rectified stimuli.

Effects of degradation type for Group 2 (impaired). Now consider the mean slopes for Group 2 separately for each degradation type. The change in log geometric mean SQMEs relative to the change in log degree of degradation was essentially the same for high-pass filtered and linearly rectified stimuli, but was significantly smaller for low-pass filtered stimuli. In other words, Group 2 required a relatively large change in degree of degradation by low-pass filtering to produce a given change in SQMEs.

Effects of degradation type for Groups 1 and 2. Inspection of all the mean slopes reveals that only the Group 1 mean for linearly rectified stimuli was greater than one. For Group 1, change in log geometric mean SQMEs was greater than change in log percent undegraded by linear rectification. As noted earlier, the between-session correlation of slopes ($n = 4$) for linearly rectified stimuli was nil ($r = 0.09$) for Group 1 and strong for Group 2. In general, change in log geometric mean SQMEs was less than change in log degree of degradation.

Further inspection of the mean slopes shows that only in the case of high-pass filtered stimuli was the Group 2 mean greater than the Group 1 mean. Recall that the difference in the means was not statistically significant and that the between-session correlation of Group 2 slopes ($n = 4$) for high-pass filtered stimuli was a weak-to-moderate one ($r = -0.52$).

For low-pass filtered and linearly rectified stimuli the change in log geometric mean SQMEs relative to the change in log degree of degradation was significantly greater for Group 1 than for Group 2. In these two instances, Group 2 required more degradation to produce a given change in SQMEs than did Group 1. In the case of high-pass filtered stimuli, Group 1 required more degradation to produce a given change in SQMEs than did Group 2. One interpretation of these findings is that changes in low frequency energy have relatively smaller effects on speech quality magnitude estimates and therefore that most listeners, particularly hearing impaired listeners, should perceive low-pass filtered signals as being of higher quality than high-pass filtered signals. This interpretation appears to be supported by recent observations (Punch and Ciechanowski, 1977; Punch and Beck, 1979; Punch and Parker, 1979; and Swartz, Walden, and Prosek, 1979) that low frequency signals are, in fact, preferred. Punch and Ciechanowski (1977) also noted better preference reliability with greater relatively low-frequency energy in the signal. Quality judgments were felt to be based mostly on low-frequency information, at least in subjects with good hearing for low frequencies.

Theoretical Factors

Perceived speech quality can be intuitively related to a variety of factors, including the amount of information transmitted and received, signal intelligibility, and a "naturalness" factor.

Information theory. Information theory (Shannon and Weaver, 1949, 1963) states that

$$I \text{ (amount of information)} = 2\ t\ w \log (S + N/N)$$

where t is signal duration, w is width of the useable frequency range, S is the maximum amplitude of the signal, and N is minimum discernible intensity difference. According to Lassman's (1964) "noise interference" model,

$$I = \frac{S/N \text{ environment}}{N_{\text{hearing aid}} + N_{\substack{\text{peripheral}\\ \text{auditory}\\ \text{system}}} + N_{\substack{\text{central}\\ \text{auditory}\\ \text{system}}}}$$

where I is intelligibility, S is signal magnitude, and N is noise in the information theory sense (i.e., anything that increases signal ambiguity). Thus, intelligibility appears to be inherently related to information transmission. Lassman (1964) was undecided about what to call his model and noted that the title almost calls for something like "a model of information flow". In any event, the most currently relevant variables in these two formulas appear to be frequency bandwidth and "noise".

For the sake of argument, assume that information theory (Shannon and Weaver, 1949, 1963) and Lassman's model are applicable to quality judgments and that speech quality increases when the amount of information transmitted and received increases. Signal presentation levels for the impaired listeners ranged from 58 to 90 dB hearing level, while the thresholds in the test ear ranged from 20 to 100 dB hearing level at 2000 Hz and from 35 dB to hearing levels beyond the intensity limits of the audiometer at 4000 Hz. Thus, the normal hearing listeners received high frequency information at higher sensation levels relative to the thresholds for high frequencies than did the impaired listeners. Although the same signal bandwidths were presented to all listeners under the various filtering conditions, the "useable" bandwidths were probably smaller for hearing impaired listeners than for normal hearing listeners due to differences in auditory systems. Although the same "noise" levels were presented to all listeners under the various filtering and linear rectification conditions, these "noise" levels may have interacted with larger "noise" levels in the impaired auditory systems than in the normal auditory systems. Different useable bandwidths and "noise" levels in the auditory system could affect the growth of perceived speech quality magnitude differently, depending upon the type of signal degradation.

As low-pass filtered signals increased from minimum (550 Hz) to maximum (4350 Hz) bandwidth, the mean geometric mean SQMEs underwent a 4.7 fold increase for Group 1 and a 1.9 fold increase for Group 2. The finding of a significant difference in the slopes obtained for Groups 1 and 2 suggests that Group 2 experienced an abnormally slow growth in perceived speech quality as a function of increasing bandwidth.

As high-pass filtered signals increased from minimum (900 Hz) to maximum (4350 Hz) bandwidth, the mean geometric mean SQMEs underwent a 2.7 fold increase for Group 1 and a 3.9 fold increase for Group 2. In this instance, it appears that Group 2 experienced an abnormally rapid growth in perceived speech quality magnitude. Recall, however, that there was no significant difference in the slopes obtained for Groups 1 and 2.

Mean geometric mean SQMEs for linearly rectified stimuli underwent a 2.8 fold increase for Group 1 and a 1.9 fold increase for Group 2, as percentages of total harmonic distortion decreased from 60% (40% undegraded) to 1% (99% undegraded). Since the observed slopes for Groups 1 and 2 were significantly different, Group 2 apparently experienced an abnormally slow growth in perceived speech quality magnitude as a function of decreasing total harmonic distortion.

In general, it can be argued that more information is received and SQMEs increase as useable bandwidth increases and "noise" level decreases. The growth rate of perceived speech quality magnitude and information received, however, appears to vary as a function of group-by-degradation type interaction. The exact relationship between SQMEs and information transmission is unknown.

Intelligibility theory. Perceived speech quality has been related to the intelligibility of the signal as demonstrated on discrimination tests (McGee, 1965).

The seven cutoff frequencies for the low-pass and high-pass filtered signals were interpolated from French and Steinberg's (1947) articulation scores for filtered CVC monosyllables. The seven cutoff frequencies were estimated from percent correct word discrimination scores ranging from 10 to 100% in 15% increments. As would be expected, SQME values increased as frequency bandwidth increased. As noted earlier, the growth of perceived speech quality magnitude as a function of low frequency bandwidth was significantly slower for Group 2 than for Group 1. This is reasonable, since the rate of change in Group 2's discrimination ability may also be slower. The growth of perceived speech quality magnitude with increases in high frequency bandwidth did not differ for Groups 1 and 2 in spite of the expectation that rate of change in discrimination ability would differ for the two groups.

Licklider (1946) noted that 50% word discrimination is roughly equivalent to 90% sentence discrimination. Interpolations from French and Steinberg's (1947) syllable articulation data suggest scores of about 55 to 100% for four of the seven degrees of degradation used for the low-pass and high-pass filtered signals in the current study. Thus, it is probably safe to assume that four of the seven degradation levels for the low-pass and high-pass filtered signals were at least 90% intelligible for normal hearing listeners. Intelligibility theory alone can account for relatively little of the change in quality related to filtering.

Percent undegraded by linear rectification varied from 40% to greater than 99% in 10% increments. Perceived speech quality magnitude increased as percent undegraded increased. The rate of increase was slower for Group 2 than for Group 1. Licklider and Held (1952) reported that their normal hearing subjects attained a 98% discrimination score on words subjected to half-wave rectification (40% THD or 60% undegraded). Since four of the seven degradation levels represent less than 40% THD, it is highly unlikely that Group 1's SQMEs were based entirely on discrimination ability. Group 2's SQMEs could be more closely related to discrimination ability since their discrimination ability should be poorer and show less improvement with changes in degree of degradation. In total, however, intelligibility theory appears to account for relatively little of the change in quality related to linear rectification.

Implications for Future Research

Speech quality magnitude estimation appears to be a manageable task for most people after minimal training. The visual training and screening task employed here appears to have served its purposes well. Since the visual data were reliable across time and groups, systematic differences in performance on the SQME tasks can be attributed to perceptual differences rather than to lack of skill in magnitude estimation. Similar training and screening procedures can be recommended for future studies involving magnitude estimation tasks.

The current study shows that log geometric mean SQMEs are positively and linearly related to log degree of degradation by low-pass filtering, high-pass filtering, and linear rectification. The study also shows that change in log geometric mean SQMEs relative to change in log degree of degradation (i.e., slope) varies as a function of degradation type and group-by-degradation type interaction. These findings have implications for additional research on the evaluation of communications systems in engineering, research on clinical practices in audiology, and further basic research to investigate how normal hearing and hearing impaired individuals perceive complex signals. When perceptual phenomena exhibit power-law behavior, some degree of prediction becomes possible.

Perception of Complex Signals

Classification of SQME continua. A ubiquitous issue in psychophysics is whether perceptual continua are of the prothetic or metathetic class. Since this issue was not resolved in the current study, a future study could be specifically designed to include functional criteria described by S. S. Stevens (1957) for distinguishing between classes of continua.

Matching perceptual experiences. An interesting question in audiology is how to determine degrees of degradation which will allow a normal hearing listener to have a perceptual experience similar to that of a hearing impaired listener. Cross-modality matching procedures (S. S. Stevens, 1975) may be helpful in answering this question. Assume that each member of a normal hearing listener group and a hearing-impaired listener group performs magnitude estimates by assigning numbers to vibrations on the finger and to the quality of speech samples varying in degree of degradation. Further assume that each of the same subjects adjusts a vibration on his finger until it matches the perceived quality of the same degraded speech samples. If the log perceptual values in each case are examined as a function of the log stimulus values, each of the functions should be linear, and the slopes of the matching function for each group should equal the ratio of the slopes for the two original functions for each group (S. S. Stevens, 1975). If the predicted slopes are verified in the matching functions, common vibration amplitudes for the two listener groups may represent speech samples that are perceived as being of equal quality. If this proves to be the case, application of cross-modality matching procedures in this manner may enable the normal hearing listener to relate more realistically to the perception of an individual who is hearing impaired.

Another interesting question relates to how much degradation of one type produces quality that is equivalent to that produced by a given degradation level of another type. Assume that a group of listeners assigns numerical magnitudes to perceived vibrations on the finger and to the perceived quality of speech samples degraded to varying degrees by low-pass filtering and high-pass filtering. Each subject would then adjust the vibration magnitude until it matches the perceived quality of the degraded speech samples. Log perceptual values should grow linearly as a function of log degree of degradation, and the slope of the matching function for each degradation type should equal the ratio of the slopes for the original vibration function and perceived speech quality function for the degradation type in question (S. S. Stevens, 1975). If the predicted slopes are observed in the matching functions, common vibration amplitudes for the two degradation types may represent speech samples that are perceived as being of equal quality.

Another application of the cross-modality matching procedure might enable one to equate quality perceived in one domain with that perceived in another. Assume that a group of subjects assigns numerical magnitude estimates to: (1) the quality of a visually projected image subjected to varying degrees of distortion, (2) the quality of speech samples degraded to varying degrees, and (3) a series of vibrations on the finger. Further assume that each subject adjusts vibration magnitude on the finger until it matches the perceived quality of the distorted visual images and the degraded speech samples. If the power law holds for each function, the slopes of the two matching functions should equal the ratios of the vibration function and the respective original quality functions (S. S. Stevens, 1975). If this happens, equal magnitudes on the two matching functions may represent equivalent quality perception in the visual and auditory modalities.

Determinants of speech quality. Perception of speech quality appears to be related to a host of interacting factors. Future studies should consider various forms of multivariate analyses to discover the determinants of quality perception and their relative importance for normal hearing and hearing impaired listeners.

Extensions of the current study. Studies similar to the current study should determine the effect of more closely defined type and degree of hearing loss and more severe hearing loss on quality perceptions. Such studies should be extended to include different degradation types, combinations of degradation types, and more degrees of degradation.

Prediction of Perceptual Experience

Scales developed from magnitude estimation and production of speech quality can be used in much the same way that the phon (Fletcher and Munson, 1933) and sone (S. S. Stevens, 1936) scales are used. For example, consider in the place of phons, SQUALs (speech quality levels). A SQUAL could be equated with a given dB level of a standard comparison signal which would be degraded to some specified degree in some specified manner (e.g., by adjusting cutoff frequencies). Equal quality contours could be developed for various dB levels of the standard comparison signal. It is likely that speech quality will vary with presentation level as well as type and degree of degradation. Now, in place of sones, consider SQUALUs (speech quality units). One SQUALU would be equal to some speech quality level (SQUAL). A speech sample of n SQUALUs would be perceived as having n times the quality of one SQUALU.

On the basis of SQUALs and SQUALUs, it is likely that predicted speech quality for various degraded speech samples could be predicted for normal hearing listeners and that a given number of SQUALUs perceived by normal hearing listeners will be equivalent to the quality perceived by a hearing impaired listener. In this latter instance, it is likely that the number of SQUALUs indicated would suggest a reduced level of perceived quality.

Evaluation of Communication Systems

Rothauser, Urbanek, and Pachl (1968) noted that evaluation and optimization criteria are needed during the design, development, and testing of speech handling and processing systems. Rothauser, Urbanek, and Pachl (1971) defined a preference unit (PU) scale based on the "Transmission Preference Units" (TPU) introduced by Munson and Karlin (1962). Rothauser *et al*. (1971) related data to the PU scale from four preference evaluation methods (the isopreference method, the category judgment method, the relative preference method, and an absolute preference judgment method). They found it impossible to recommend any of the four as a single best method for all situations. Thus, speech quality magnitude estimation and production procedures may prove to be helpful in determining evaluation and optimization criteria for engineering purposes. Such an application would probably require the development of some sort of quality scale (e.g., the SQUAL and SQUALU scales).

Clinical Application

Description. Log SQME-log degree of degradation functions may be useful in describing clinical problems. The slopes of such functions have been shown to vary as a function of degradation type and the interaction of degradation type with hearing loss, suggesting the occurrence of different perceptual events. As more is learned about the processing of complex signals, measures like the SQUALs and SQUALUs described above may become useful in describing perceptual difficulty.

Diagnosis. Audiological diagnosis typically provides information which contributes to determining site of lesion in the auditory system. The most difficult diagnostic problems are probably those which require differentiation among the cochlea, the eighth nerve, and the central auditory nervous system as the site of lesion. Direct magnitude scaling procedures have been used on experimental bases to identify recruitment due to cochlear pathology. Using a cross-modality matching procedure to study the loudness growth of pure tones in patients with unilateral conductive hearing loss and patients with unilateral sensorineural hearing loss, Thalmann (1965) found that loudness balances of the right ear against the left ear were predicted by vibration matches on the fingers. S. S. Stevens (1966) indicated that abnormal loudness growth can best be described by two straight lines in a log-log plot. He postulated that whenever the exponent (slope) of a sensory function is altered by pathology or some other circumstance, a power transformation occurs. In sensorineural hearing loss due to cochlear pathology, the power transformation is related to recruitment. In a similar vein, research on direct magnitude scaling of speech quality by sensorineurally impaired listeners may lead to even finer discrimination of hearing loss on the basis of slopes or power transformations revealed in a log-log plot. It is generally accepted that speech signals are very helpful as test materials for distinguishing among disorders of the cochlea, eighth nerve, and central auditory nervous system (Katz, 1962; Bocca and Calearo, 1963; and Jerger and Jerger, 1971, 1975).

Audiological diagnosis typically provides information about type and degree of hearing loss or deficit, but often provides very little information about hearing handicap (Oyer and Frankmann, 1975). It may be that communicative effectiveness in everyday life is closely related to the magnitude of the overall speech quality perceived by a listener. If a suitably strong relation exists, clinical measures of quality perception may facilitate assessment of oral-aural communicative integrity.

Prognosis and progress. The slopes of the log-log functions in the present study differed as a function of speech degradation type and listener group-by-degradation type interaction. If perceived speech quality magnitude can be improved by therapy, it may be that speech quality magnitude functions obtained at different points in time can serve as measures of progress in aural rehabilitation. Certain slope values might become valuable as predictors of success in improving perceptual skills.

Where physical stimulus characteristics are determined by different hearing aids, the relative speech quality magnitudes assigned to signals transduced by different aids may be helpful in hearing aid selection procedures. Previous investigators have indicated that quality judgments are related to the electroacoustic characteristics of hearing aids (Jeffers, 1960; Zerlin, 1962; Witter and Goldstein, 1971; Smaldino, 1974; Punch and Ciechanowski, 1977; and Chial and Daniel, 1977) and that quality judgments may be more sensitive to electroacoustic differences than are traditionally used word discrimination tests (Jeffers, 1960 and Zerlin, 1962). Quality judgments may represent a good predictor of success in using amplification.

CHAPTER V

SUMMARY AND CONCLUSIONS

Introduction

Background

Traditional word discrimination tests do not adequately predict listener performance in everyday life and therefore do not measure handicap (Chial and Hayes, 1974; Oyer and Frankmann, 1975; Millin, 1975; and Berger, 1978). Speech quality judgments have sometimes differentiated among hearing aids when word discrimination tests did not (Jeffers, 1960; Zerlin, 1962). Thus, judgments of speech quality may be more closely related to everyday communicative effectiveness than are word discrimination scores. Although there appears to be some interest in the use of quality judgments in clinical audiology, there has been insufficient research with hearing impaired subjects to warrant their routine use. Furthermore, most of the available research has been limited to the method of paired comparisons. Although direct magnitude estimation procedures appear to offer some advantages over paired comparisons, no one has studied this method to determine whether Stevens' power law applies to the scaling of speech quality and whether there are differences in speech quality magnitude estimate-degree of degradation functions as a function of hearing status and signal degradation type.

Purpose

This study was designed to obtain speech quality magnitude estimates (SQMEs) from normal hearing and sensorineurally impaired hearing listeners as a function of seven degrees of degradation by low-pass filtering, high-pass filtering, and linear rectification. The log geometric mean SQME-log degree of degradation functions were analyzed:

1. to determine whether log geometric mean SQMEs differed as a function of changes in log degree of degradation for each listener group under each degradation type;

2. to identify the lowest order equation required to provide a satisfactory fit to the log-log functions; and

3. to determine whether differences existed among the slopes of the log-log functions as a function of listener group, degradation type, or group-by-degradation type interaction.

Experimental Design

Subjects

There were two groups of subjects. Group 1 consisted of 12 normal hearing listeners with a mean age of 23.58 years, a mean two-frequency average hearing threshold level of 0.42 dB in the test ear, and a mean speech discrimination score of 99.5 percent in the test ear. Group 2 consisted of 12 sensorineurally impaired hearing listeners with a mean age of 32.58 years, a mean two-frequency average hearing threshold of 33.50 dB in the test ear, and a mean speech discrimination score of 80.17 percent in the test ear. For the most part, the Group 2 hearing losses were most pronounced for high frequencies. Their hearing test results were consistent with a cochlear site of lesion.

Stimuli

Visual training and screening stimuli. The visual stimuli consisted of three sets of seven pairs of geometric forms located side by side on 2" x 2" slides. The left member of each pair was always the standard stimulus, whose size did not change; the right member of each pair was always one of the seven comparison stimuli which differed in size. The first set of seven pairs consisted of squares, while the last two sets of seven pairs consisted of circles. Each set represented a trial.

Auditory stimuli. The auditory stimuli were audio recordings of six 10-second connected speech samples from a junior high history text, which were, in their undegraded form, approximately equal in listening difficulty. The final test tapes consisted of degraded versions of these samples recorded in pairs. Each pair consisted of a standard stimulus followed by one of seven comparison stimuli. Each comparison stimulus represented one of seven degradation levels, while the standard stimulus always represented the middle (fourth) of the seven degrees of degradation. The types of degradation were low-pass filtering, high-pass filtering, and linear rectification. Degrees of degradation were represented by: (1) bandwidths of 550, 950, 1300, 1650, 1950, 2950, and 4350 Hz for low-pass filtering; (2) bandwidths of 900, 1400, 1600, 2100, 2550, 3000, and 4350 Hz for high-pass filtering; and (3) percent undegraded values of 40%, 50%, 60%, 70%, 80%, 90%, and 99% for total harmonic distortion due to linear rectification.

Procedures

All subjects underwent: (1) a hearing screening, (2) visual magnitude estimation training and screening, (3) SQME training, and (4) the SQME experiment in that order. Four subjects in each listener group returned to repeat the SQME training and the SQME experiment at a later date.

Hearing screening. The hearing screening consisted of a brief history, pure tone air conduction testing, reflex decay or tone decay testing, word discrimination testing in the test ear, and impedance testing.

Visual magnitude estimation training and screening. Each subject was presented with three sets of randomly ordered pairs of visual stimuli, which were accompanied by audio taped instructions. The first set consisted of squares which served as training stimuli; the second and third sets consisted of circles which provided a test-retest reliability check. Each subject assigned a numerical value to the magnitude of the standard stimulus in each set and then estimated the magnitudes of the comparison stimuli relative to that of the standard. The visual task served as a procedural model for the SQME tasks.

SQME training. Twenty-one tape recorded practice items (seven degrees of each of the three degradation types) preceded the SQME experiment. The same two passages, one for the standard stimulus and one for the comparison stimulus, were used for all practice items. All stimuli of a single degradation type were presented in succession, with degradation levels being presented in random order. The order of presentation for degradation types was counterbalanced across pairs of listeners. Listeners assigned SQMEs to the auditory standard and comparison stimuli in the same manner used to make visual magnitude estimates in the visual tasks.

SQME experiment. SQMEs were obtained on 42 experimental items (two seven-item trials of each of the three degradation types) following the same procedure used in the visual and SQME training tasks. In the experimental tasks, however, there were four, rather than two, stimulus passages. Two served as standard stimuli, and two served as comparison stimuli. The same two passages were always presented together, so that there were, in effect, only two different pairings of the four stimuli. The order of these pairs of passages was counterbalanced across trials. Just as in the SQME training task, the order of degradation levels was randomized within each trial, and the order of degradation types was counterbalanced across pairs of listeners.

Dependent Variables

Two dependent variables were derived from the visual magnitude estimates and the SQMEs. The visual magnitude estimates provided the basis for (1) log visual magnitude estimates and (2) the slopes of the lines of best fit which relate log visual magnitude estimates to log circle size. The SQMEs provided the basis for (1) log geometric mean SQMEs within subjects and across trials and (2) the slopes of the lines of best fit relating log geometric mean SQMEs to log degree of degradation.

Findings

Findings of the study include the following.

1. Log visual magnitude estimates appear, as indicated by S. S. Stevens (1975), to be linearly related to log circle area and therefore, to represent a power function.

2. The mean slopes for the log visual magnitude - log area functions produced by a group of normal hearing listeners and a group of sensorineurally impaired hearing listeners are in agreement with the slope reported by S. S. Stevens (1975) for similar data.

3. Log geometric mean SQMEs (across trials) are positively and linearly related to log bandwidth of low-pass and high-pass filtered stimuli and to log percent undegraded by linear rectification, indicating that a power function applies in each case.

4. Slopes of the log-log functions for Group 1 differed from those for Group 2 under low-pass filtering and linear rectification, but not under high-pass filtering.

5. For Group 1 the slopes of the log-log functions for low-pass and high-pass filtered stimuli did not differ from each other, but did differ from the slopes of the log-log functions for linearly rectified stimuli.

6. For Group 2 the slopes of the log-log functions for high-pass filtered and linearly rectified stimuli did not differ from each other, but did differ from the slopes of the log-log functions for low-pass filtered stimuli.

Conclusions

The results of this study seem to provide the basis for the following tentative conclusions.

1. Excellent reliability of listener performance on the visual task within subjects, across groups, and across time suggests that systematic differences in performance on the SQME tasks are probably due to perceptual differences.

2. Tasks similar to the visual training and screening task used in this study appear to be practical for use in future studies involving the method of magnitude estimation.

3. The slopes of the log geometric mean SQME - log degree of degradation functions differ as a function of group-by-degradation type interaction, suggesting systematic perceptual differences.

4. Estimates of poor between-session reliability of slopes suggested that Group 1 may have had perceptual difficulty with linearly rectified stimuli, while Group 2 may have had perceptual difficulty with high-pass filtered stimuli.

5. Factors other than discrimination ability and information transmission and reception appear to play a major role in the perception of speech quality.

6. The findings of this study were sufficiently encouraging to warrant additional research on the evaluation of communications systems, clinical practices in audiology, and how normal hearing and hearing impaired individuals process complex signals.

APPENDICES

APPENDIX A

TABULAR SUMMARY OF AGE AND AUDIOMETRIC DATA FOR INDIVIDUAL SUBJECTS AND GROUPS

The two tables which follow summarize the age and audiometric data for individual subjects and groups. Table A-1 summarizes the ages, two-frequency average hearing threshold levels, test ear discrimination scores, and the means and standard deviations for all normal hearing subjects (Group 1) and all sensorineurally impaired hearing subjects (Group 2). Table A-2 lists the pure tone thresholds and median thresholds as a function of ear and test frequency for the hearing impaired subjects.

Table A-1. A summary of ages, two-frequency average thresholds, test ear discrimination scores, and the means and standard deviations for all subjects.

	Subject	Age Yrs.	2-Frequency Average Threshold R dB	2-Frequency Average Threshold L dB	% Discrimination Score
Group 1	(Normal)				
	1	22	0[a]	0	100
	2	24	0	0	100
	3	23	0	0	100
	4	24	0	0	100
	5	23	0	0	100
	6	22	0	0	98
	7	27	0	0	98
	8	28	5	5	100
	9	22	0	0	100
	10	22	0	0	98
	11	23	0	0	100
	12	23	0	0	100
$\bar{X}$		23.58	0.42	0.42	99.5
S.D.		1.98	1.44	1.44	0.9
Group 2	(Impaired)				
	13	36	48	45[b]	90
	14	49	25	20[b]	84
	15	21	33	28[b]	90
	16	35	35[b]	37	70
	17	48	18	18[b]	70
	18	20	48	40[b]	88
	19	19	60	50[b]	60
	20	18	35[b]	45	90
	21	46	35	25[b]	74
	22	37	48[b]	50	90
	23	30	20	15[b]	88
	24	32	43[b]	73	68
$\bar{X}$		32.58	37.33	37.17	80.17
S.D.		11.37	12.61	16.87	10.97

[a]The right ear was the test ear for all subjects in Group 1.

[b]Test ear thresholds for Group 2: $\bar{X} = 33.50$, S.D. $= 12.13$.

Table A-2. Pure tone thresholds (dB) and median thresholds as a function of ear (R, L) and frequency (Hz) for the hearing impaired subjects.

	250 Hz		500 Hz		1000 Hz		2000 Hz		4000 Hz		6000 Hz	
Subject	R	L	R	L	R	L	R	L	R	L	R	L
13	20	20	35	35	60	55	65	55	55	50	50	65
14	30	20	25	15	25	25	55	55	75	75	85	NR[a]
15	25	25	25	15	40	40	40	45	50	45	60	65
16	10	10	15	15	50	60	100	NR[a]	NR[a]	NR[a]	NR[a]	NR[a]
17	5	0	15	5	20	40	65	50	80	75	95	70
18	35	35	40	35	55	45	55	50	65	70	65	65
19	30	10	45	30	75	75	105	85	NR[a]	NR[a]	NR[a]	NR[a]
20	25	30	30	40	40	50	45	55	60	60	55	55
21	20	15	20	15	45	35	55	55	55	55	55	55
22	40	45	50	50	55	55	45	50	35	50	40	40
23	15	20	15	15	25	15	45	20	90	55	80	60
24	0	35	30	65	75	80	55	85	40	60	15	70
Median	22.5	20.0	27.5	22.5	47.5	47.5	55.0	55.0	62.5	60.0	62.5	65.0
Test Ear Median	20.0		22.5		42.5		52.5		57.5		65.0	

[a] NR: No response at intensity limits of the audiometer.

APPENDIX B

TRANSCRIPT OF SIX STIMULUS PASSAGES

<u>Practice Standard B</u>

Balboa named his discovery the South Sea because it lay directly south of where he started his march. It was not until after Magellan's voyage that the sea was called Pacific, the name we use today (Wilder *et al.*, 1954, p. 62).

<u>Practice Comparison B</u>

After several years spent in preparation, Pizarro set off on his great adventure. He landed safely on the coast of Peru, where he remained for some time 'sizing up' the situation (Wilder *et al.*, 1954, p. 66).

<u>Experimental Standard A</u>

Two months later the weary Spaniards stood looking in amazement upon the Aztec capital. The city was built on islands in the center of a large lake, and was connected with the mainland by three roads or causeways (Wilder *et al.*, 1954, p. 64).

<u>Experimental Comparison A</u>

One of the early settlers, named John Rolfe, learned how to produce fine tobacco. Smoking was becoming popular in England, so the Jamestown colonists found it easy to sell all the tobacco that could be grown (Wilder *et al.*, 1954, p. 87).

<u>Experimental Standard C</u>

The bold explorers who searched this land did not find the waterway they were seeking, but they accomplished something more important. They turned the attention of Europe away from Asia to the New World itself (Wilder *et al.*, 1954, p. 53).

<u>Experimental Comparison C</u>

Every kind of disaster happened to the expedition - storms, sickness, death, mutiny, desertion. But at last the men who remained alive anchored once more in a Spanish harbor (Wilder *et al.*, 1954, p. 43).

Source: Wilder, H. B., Ludhum, R. P. and Brown, H. M. *This is America's Story*. Boston: Houghton, Mifflin Company (1954).

APPENDIX C

TABULAR DESCRIPTION OF THE SUBMASTER RECORDINGS, RECORDINGS PRODUCED BY THE COMPUTER SYSTEM, AND THE FINAL TEST TAPES

The three tables which follow describe the makeup of the three sets of recordings generated from the master tape. Table C-1 summarizes the nine submaster recordings. Table C-2 describes the interim recordings produced from the submaster recordings by the computer system. Table C-3 describes the final test tapes which were spliced from sections of the interim tapes produced by the computer system.

Table C-1. Summary of the nine submaster recordings.

Submaster Cassette No.	Contents	Degradation Type	Degradation Level
1	Comparison Passages A,B,C	L-PF H-PF LR	1 1 1
2	Comparison Passages A,B,C	L-PF H-PF LR	2 2 2
3	Comparison Passages A,B,C	L-PF H-PF LR	3 3 3
4	Comparison Passages A,B,C	L-PF H-PF LR	4 4 4
5	Comparison Passages A,B,C	L-PF H-PF LR	5 5 5
6	Comparison Passages A,B,C	L-PF H-PF LR	6 6 6
7	Comparison Passages A,B,C	L-PF H-PF LR	7 7 7
8	Standard Passages A,B,C	L-PF H-PF LR	4 4 4
9	Labels (i.e., Trial 1, standard, item, one, etc.)	None	None

Table C-2. Crossbreak matrix showing the makeup of the recordings produced by the computer system.

Interim Tape No.	Three Degradation Types[a]	Pairing of Six Stimulus Passages[b]	Number of Random Orders
1	L-PF	PSB & PCB_1-PCB_7	6
2	H-PF	PSB & PCB_1-PCB_7	6
3	LR	PSB & PCB_1-PCB_7	6
4	L-PF	ESA & ECA_1-ECA_7	6
5	L-PF	ESC & ECC_1-ECC_7	6
6	H-PF	ESA & ECA_1-ECA_7	6
7	H-PF	ESC & ECC_1-ECC_7	6
8	LR	ESA & ECA_1-ECA_7	6
9	LR	ESC & ECC_1-ECC_7	6

[a]L-PF: low-pass filtering
H-PF: high-pass filtering
LR : linear rectification

[b]PSB: practice standard passage B
PCB_1-PCB_7: 7 degradation levels of practice comparison passage B
ESA: experimental standard passage A
ESC: experimental standard passage C
ECA_1-ECA_7: 7 degradation levels of experimental comparison passage A
ECC_1-ECC_7: 7 degradation levels of experimental comparison passage C

Table C-3. Crossbreak matrix: Presentation orders for degradation types, passages, and random orders for comparison degradation levels.

Subjects	Trial	Three Degradation Types[a]	Pairings of Six Stimulus Passages[b]	Order of Seven Comparison Degradation Levels
1 & 2	1	LR	PSB & PCB_1-PCB_7	2135764
(13 & 14)	2	L-PF	PSB & PCB_1-PCB_7	3246517
	3	H-PF	PSB & PCB_1-PCB_7	4126375
	1	L-PF	ESA & ECA_1-ECA_7	7631524
	2	L-PF	ESC & ECC_1-ECC_7	6274513
	1	H-PF	ESA & ECA_1-ECA_7	5374621
	2	H-PF	ESC & ECC_1-ECC_7	1426375
	1	LR	ESA & ECA_1-ECA_7	7613542
	2	LR	ESC & ECC_1-ECC_7	1743652
3 & 4	1	LR	PSB & PCB_1-PCB_7	2653741
(15 & 16)	2	H-PF	PSB & PCB_1-PCB_7	6543721
	3	L-PF	PSB & PCB_1-PCB_7	2351764
	1	L-PF	ESC & ECC_1-ECC_7	5764231
	2	L-PF	ESA & ECA_1-ECA_7	4352761
	1	LR	ESC & ECC_1-ECC_7	4267153
	2	LR	ESA & ECA_1-ECA_7	7631524
	1	H-PF	ESC & ECC_1-ECC_7	7325416
	2	H-PF	ESA & ECA_1-ECA_7	7651342
5 & 6	1	H-PF	PSB & PCB_1-PCB_7	2453671
(17 & 18)	2	LR	PSB & PCB_1-PCB_7	2574316
	3	L-PF	PSB & PCB_1-PCB_7	1463275
	1	H-PF	ESA & ECA_1-ECA_7	2475163
	2	H-PF	ESC & ECC_1-ECC_7	4657123
	1	L-PF	ESA & ECA_1-ECA_7	3251647
	2	L-PF	ESC & ECC_1-ECC_7	1426375
	1	LR	ESA & ECA_1-ECA_7	2143657
	2	LR	ESC & ECC_1-ECC_7	4723156
7 & 8	1	H-PF	PSB & PCB_1-PCB_7	3726154
(19 & 20)	2	L-PF	PSB & PCB_1-PCB_7	6245137
	3	LR	PSB & PCB_1-PCB_7	4316572
	1	H-PF	ESC & ECC_1-ECC_7	3527641
	2	H-PF	ESA & ECA_1-ECA_7	1547236
	1	LR	ESC & ECC_1-ECC_7	6512374
	2	LR	ESA & ECA_1-ECA_7	3152647
	1	L-PF	ESC & ECC_1-ECC_7	1435276
	2	L-PF	ESA & ECA_1-ECA_7	6574132

Table C-3 (cont'd).

Subjects	Trial	Three Degradation Types[a]	Pairings of Six Stimulus Passages[b]	Order of Seven Comparison Degradation Levels
9 & 10	1	L-PF	PSB & PCB_1-PCB_7	4271635
(21 & 22)	2	LR	PSB & PCB_1-PCB_7	7326451
	3	H-PF	PSB & PCB_1-PCB_7	2134756
	1	LR	ESA & ECA_1-ECA_7	5463712
	2	LR	ESC & ECC_1-ECC_7	7123645
	1	H-PF	ESA & ECA_1-ECA_7	7352416
	2	H-PF	ESC & ECC_1-ECC_7	6274135
	1	L-PF	ESA & ECA_1-ECA_7	5176324
	2	L-PF	ESC & ECC_1-ECC_7	4716235
11 & 12	1	L-PF	PSB & PCB_1-PCB_7	3742156
(23 & 24)	2	H-PF	PSB & PCB_1-PCB_7	3267451
	3	LR	PSB & PCB_1-PCB_7	2156437
	1	LR	ESC & ECC_1-ECC_7	1357462
	2	LR	ESA & ECA_1-ECA_7	5164732
	1	L-PF	ESC & ECC_1-ECC_7	3657241
	2	L-PF	ESA & ECA_1-ECA_7	5412637
	1	H-PF	ESC & ECC_1-ECC_7	6274315
	2	H-PF	ESA & ECA_1-ECA_7	4531726

[a]L-PF: low-pass filtering
H-PF: high-pass filtering
LR : linear rectification

[b]PSB: practice standard passage B
PCB_1-PCB_7: 7 degradation levels of practice comparison passage B
ESA: experimental standard passage A
ESC: experimental standard passage C
ECA_1-ECA_7: 7 degradation levels of experimental comparison passage A
ECC_1-ECC_7: 7 degradation levels of experimental comparison passage C

APPENDIX D

GLADC: COMPUTER PROGRAM USED FOR ANALOG-TO-DIGITAL CONVERSIONS

```
  FILE  GLADC.MAC       19-JUL-78              PAGE    1


.TITLE  MLAT
.GLOBL  MOF$MS,MUF$MS,CKF$,MON$SM


 .MCALL  ..V2..,..REGDEF,..STRUCT,.READW,.WRITW,.PRINT
 .MCALL  .TTYOUT,[illegible],[illegible],.CLOSE,.EXIT,[illegible],[illegible]
         ..V2..
         .REGDEF
         .STRUCT


RDPTR:   .WORD    0,0        ;READ POINTER
RDLOW:   .WORD    0,0        ;LOW BUFFER
RDHIGH:  .WORD    0,0        ;HIGH BUFFER
RDSIZE:  .WORD    0          ;READ FILE SIZE (BLOCKS)
BUF1:    .BLKW    3072.
BUF2:    .BLKW    3072.

READWD:  MOV      R1,-(SP)
         MOV      R2,-(SP)
         MOV      RDPTR,R1
         CLRB     R1
         BISB     RDPTR+2,R1
         SWAB     R1
         IF <<R1 GE RDSIZE> OR <RDPTR+2 LT #0>>
         CLR      R0
         MOV      (SP)+,R2
         MOV      (SP)+,R1
         SEC
         ADC      RDPTR
         ADC      RDPTR+2
         RTS      PC
         END
         MOV      RDPTR+2,R1          ;HIGH PART IN R1
         MOV      RDPTR,R2            ;LOW PART IN R2
         CLR      -(SP)     ;DATA NOT IN CORE FLAG.
         IF <R1 LT RDLOW+2>
         DEC      (SP)
         END
         IF <<R1 EQ RDLOW+2> AND <R2 LO RDLOW>>
         DEC      (SP)
         END
         IF <R1 GT RDHIGH+2>
         DEC      (SP)
         END
         IF <<R1 EQ RDHIGH+2> AND <R2 HIS RDHIGH>>
         DEC      (SP)
         END
         IF <(SP)+ NE #0>
         MOV      RDPTR,R2
```

```
        CLRB    R2
        BISB    RDPTR+2,R2
        SWAB    R2
        MOV     RDSIZE,R1
        SUB     #12.,R1
        IF  R2 LE R1
        .READW  #AREA,#0,#BUF1,#3072.,R2
        SWAB    R2
        CLR     R1
        BISB    R2,R1
        CLRB    R2
        MOV     R2,RDLOW
        MOV     R1,RDLOW+2
        ADD     #3072.,R2
        ADC     R1
        MOV     R2,RDHIGH
        MOV     R1,RDHIGH+2
        ELSE
        MOV     RDSIZE,R1
        SUB     R2,R1
        SWAB    R1
        .READW  #AREA,#0,#BUF1,R1,R2
        MOV     R1,R0
        SWAB    R2
        CLR     R1
        BISB    R2,R1
        CLRB    R2
        MOV     R2,RDLOW
        MOV     R1,RDLOW+2
        ADD     R0,R2
        ADC     R1
        MOV     R2,RDHIGH
        MOV     R1,RDHIGH+2
        END
        END
        MOV     RDPTR,R2                ;LOW PART IN R2
        SUB     RDLOW,R2
        ASL     R2
        ADD     #BUF1,R2
        MOV     (R2),R0
        SEC
        ADC     RDPTR
        ADC     RDPTR+2
        MOV     (SP)+,R2
        MOV     (SP)+,R1
        RTS     PC


WTPTR:  .WORD   0,0             ;COUNTER
WTLOW:  .WORD   0,0             ;WRITE BUFFER LOW LIMIT
WTHIGH: .WORD   0,0             ;WRITE BUFFER HIGH LIMIT
```

```
WTSIZE: .WORD   0               ;WRITE FILE SIZE (BLOCKS)

WRITWD:
        .PSHREG
        CLR     -(SP)
        IF ((WTPTR+2 LT WTLOW+2) OR (WTPTR+2 GT WTHIGH+2))
        DEC     (SP)
        END
        IF ((WTPTR+2 EQ WTLOW+2) AND (WTPTR LO WTLOW))
        DEC     (SP)
        END
        IF ((WTPTR+2 EQ WTHIGH+2) AND (WTPTR HIS WTHIGH))
        DEC     (SP)
        END

        IF (SP)+ NE #0
        MOV     R0,-(SP)
        JSR     PC,WRITES
        MOV     (SP)+,R0
        END
        MOV     WTHIGH,R5
        MOV     WTHIGH+2,R4
        MOV     WTLOW,R3
        MOV     WTLOW+2,R2
        IF ((R2 LT R4) OR ((R2 EQ R4) AND (R3 LO R5)))
        MOV     WTPTR,R1
        SUB     WTLOW,R1
        ASL     R1
        MOV     R0,BUFI(R1)
        END
        SEC
        ADC     WTPTR
        ADC     WTPTR+2
        .POPREG
        RTS     PC



WRITES:
        MOV     WTHIGH,R5
        MOV     WTHIGH+2,R4
        MOV     WTLOW,R3
        MOV     WTLOW+2,R2
        SUB     R3,R5
        SBC     R4
        SUB     R2,R4
        IF ((R4 EQ #0) AND (R5 NE #0))
        CLRB    R3
        BISB    R2,R3
        SWAB    R3
        .WRITW  #AREA,#1,#BUFI,R5,R3
```

```
          END
          MOV     WTPTR,WTLOW
          MOV     WTPTR+2,WTLOW+2
          CLRB    WTLOW
          MOV     WTLOW,WTHIGH
          MOV     WTLOW+2,WTHIGH+2
          ADD     #1000,WTHIGH
          ADC     WTHIGH+2
          IF (WTLOW+2 LT #0)
          CLR     WTLOW
          CLR     WTLOW+2
          END
          MOV     WTSIZE,R1
          CLR     R0
          SWAB    R1
          BISB    R1,R0
          CLRB    R1
          IF ((R0 LT WTHIGH+2) OR ((R0 EQ WTHIGH+2) AND (R1 LO WT
*         HIGH)))
          MOV     R1,WTHIGH
          MOV     R0,WTHIGH+2
          END
          MOV     WTHIGH,R5
          MOV     WTHIGH+2,R4
          MOV     WTLOW,R3
          MOV     WTLOW+2,R2
          SUB     R3,R5
          SBC     R4
          SUB     R2,R4
          IF ((R4 EQ #0) AND (R5 NE #0))
          CLRB    R3
          BISB    R2,R3
          SWAB    R3
          .READW  #AREA,#1,#BUF2,R5,R3
          END
          RTS     PC

FLUSH:    MOV     R0,-(SP)
          MOV     R1,-(SP)
          MOV     R2,-(SP)
          MOV     WTPTR,R0
          SUB     WTLOW,R0
          MOV     R0,R1
          ADD     #377,R1
          CLRB    R1
          MOV     R1,R2
          SUB     R0,R1
          ASL     R0
          ADD     #BUF2,R0
          LOOP    R1
          CLR     (R0)+
          END
```

```
        MOV     [illegible]
        CLRB    R1
        BISB    [illegible],R1
        SWAB    R1
        .WRITW  [illegible]
        MOV     [illegible]
        MOV     [illegible]
        MOV     [illegible]
        RTS     PC

[illegible]
        LOOP
        .INPUT  #INBUF
        .FINDS  #AREA,#[illegible],#LNBUF-1
        MOV     #-1,YESFLG
        DTNEIF  R0 NE #0
        .FINDS  #AREA,#NO,#[illegible]-1
        CLR     YESFLG
        DONEIF  R0 NE #0
        .PRINT  #ERR1
        END
        RTS     PC

[illegible] .PRINT  #TCRMES
        .INPUT
        RTS     PC

ERR1:   .ASCII  /PLEASE TYPE YES OR NO? /
        .BYTE   200
TCRMES: .ASCII  /TYPE CARRIAGE RETURN TO PROCEED./
        .BYTE   200
DUMFIL: .ASCIZ  /[illegible]/
        .ASCII  / /
LNBUF:  .BLKB   81.
MES1:   .ASCII  /COPY TO NEW FILE? /
        .BYTE   200
MES2:   .ASCII  /INPUT NAME OF NEW FILE TO BE COPIED TO. /
        .BYTE   200
MES3:   .ASCII  /INPUT START TIME, STOP TIME? /
        .BYTE   200
YES:    .ASCIZ  /YES/
NO:     .ASCIZ  /NO/
        .EVEN
AREA:   .BLKW   10.
YESFLG: .WORD   0
[illegible] .WORD   0,0
[illegible] .WORD   0,0
[illegible] .FLT2   0
[illegible] .FLT2   0
TENTHO: .FLT2   10000.
```

```
REPEAT: .RADEJ  <#NO>

BEGIN:
        .LONGJ
        .OPEN   (AREA+4),,#CHANFL
        MOV     R0,ROSIZE
        CLR     ROLOW
        CLR     ROLOW+2
        CLR     ROHIGH
        CLR     ROHIGH+2
        CLR     R1
        SWAB    R0
        BISB    R0,R1
        CLRB    R0
        MOV     R0,X2
        MOV     R1,X2+2
        CLR     X1
        CLR     X1+2
        .RNADC  #BUF1,#0,#3,#*,ROUT
ASKIN:  .PRINT  #MES1
        JSR     PC,YESNO
        IF (YESFLG NE #0) THEN COPY
        .PRINT  #MES3
        .INPUT  #INBUF
        .STRFLT #AREA,#INBUF,#X1+#2
        .POLISH
        MOF$MS
        X1
        MUF$MS
        TENTHO
        CNF$
        MON$SM
        X1
        MOF$MS
        X2
        MUF$MS
        TENTHO
        CNF$
        MON$SM
        X2
        .UNPOL

        IF (X1+2 LT #0)
        CLR     X1
        CLR     X1+2
        END

        MOV     ROSIZE,R1
        CLR     R0
        SWAB    R1
        BISB    R1,R0
```

```
        CLRB    R1
        IF <X2+2 GT R0> OR <<X2+2 EQ R0> AND <X2 HI R1>>
        MOV     R0,X2+2
        MOV     R1,X2
        END

        MOV     X1,R3
        MOV     X1+2,R2
        SWAB    R3
        SWAB    R2
        CLRB    R2
        BISB    R3,R2
        CLRB    R3
        SWAB    R3

        MOV     X2,R5
        MOV     X2+2,R4
        SWAB    R5
        SWAB    R4
        CLRB    R4
        BISB    R5,R4
        CLRB    R5
        SWAB    R5
        .RKDAC  #BUF1,#0,#3,#M/ROUT,R2,R4,R3,R5
        JMP     AGAIN

COPY:   .PRINT  #MES2
        .NEWFIL #AREA,#1,#-1,#DEFEXT
        MOV     R0,WTSIZE
        CLR     WTLOW
        CLR     WTLOW+2
        CLR     WTHIGH
        CLR     WTHIGH+2
        CLR     WTPTR
        CLR     WTPTR+2
        MOV     X1,RDPTR
        MOV     X1+2,RDPTR+2
        MOV     X2,R5
        MOV     X2+2,R4
        SUB     X1,R5
        SBC     R4
        SUB     X1+2,R4
        MOV     RDPTR,R3
        MOV     RDPTR+2,R2
        LOOP
        SEC
        SBC     R5
        SBC     R4
        DONEIF  <R4 LT #0>
        JSR     PC,READWD
```

```
        JSR     PC,WRITWD
        EMD
        JSR     PC,FLUSH
        .CLOSE  #1
        .CLOSE  #2
        .EXIT


        .END    BEGIN
```

APPENDIX E

RPLAY: COMPUTER PROGRAM USED TO CONVERT DIGITAL SOUND FILES TO ANALOG FORM AND PLAY THEM IN RANDOM ORDER

```
  FILE  RPLAY.MAC      1?-JUL-7?               PAGE   1


.TITLE  RPLAY
.GLOBL  [illegible]
;
;
;
;
;
;       COPYRIGHT (C)  1977, BY MICHAEL J. STOUFFER, AUDIOLOGY
;       AND   SPEECH   SCIENCE   DEPARTMENT,   MICHIGAN   STATE
;       UNIVERSITY.
;
;
;
;       THIS SOFTWARE IS FURNISHED FOR USE ONLY ON  A  DIGITAL
;       EQUIPMENT CORPORATION PDP 11/40 RUNNING UNDER AN RT-11
;       V02C F/B MONITOR, AND MAY  BE  COPIED  ONLY  WITH  THE
;       INCLUSION   OF   THE  ABOVE  COPYRIGHT  NOTICE.   THIS
;       SOFTWARE, OR ANY OTHER  COPIES  THEREOF,  MAY  NOT  BE
;       PROVIDED  OR  OTHERWISE  MADE  AVAILABLE  TO ANY OTHER
;       PERSON EXCEPT FOR USE ON SUCH SYSTEM AND  TO  ONE  WHO
;       AGREES TO THESE LICENSE TERMS.  TITLE TO AND OWNERSHIP
;       OF THE SOFTWARE SHALL  AT  ALL  TIMES  REMAIN  IN  THE
;       AUDIOLOGY AND SPEECH SCIENCE DEPARTMENT.
;
;
;
;       THE INFORMATION IN THIS DOCUMENT IS SUBJECT TO  CHANGE
;       WITHOUT  NOTICE  AND  SHOULD  NOT  BE  CONSTRUED  AS A
;       COMMITTMENT BY EITHER THE AUDIOLOGY AND SPEECH SCIENCE
;       DEPARTMENT OR BY THE AUTHOR.
;
;
;
;
;       NO RESPONSIBILITY IS ASSUMED BY EITHER THE  AUTHOR  OR
;       THE  AUDIOLOGY  AND  SPEECH SCIENCE DEPARTMENT FOR THE
;       USE OR RELIABILITY OF THIS SOFTWARE ON OTHER THAN  THE
;       AFORE MENTIONED SYSTEM.
;
;
;
```

```
;
;       FILES REQUIRED FOR VOICE LABELING;
;               ITEM.SND,ONE.SND,TWO.SND,THREE.SND
;               FOUR.SND,FIVE.SND,SIX.SND,SEVEN.SND
;
;       COMPILE-TIME LIBRARIES REQUIRED:
;               ASCLIB,FORLIB
;
;
;       ASSEMBLE THIS PROGRAM PRECEDED BY THE
;               FILE ASCMAC.MAC
;       THIS ROUTINE USES THE STRUCTURE MACROS.
;
;
;       THIS PROGRAM IS A GENERAL-PURPOSE SOUND-FILE
;       OUTPUT  ROUTINE  USED  FOR PLAYING A LIST OF
;       SOUND FILES  IN  A  RANDOM  ORDER.   SEVERAL
;       OPTIONS   MAKE   THE  USE  OF  THIS  PROGRAM
;       ESPECIALLY ATTRACTIVE.  THE  VOICE-LABELING
;       OPTION  ALLOWS  THE  USER  TO  ATTACH SPOKEN
;       LABELS  TO  THE  OUTPUT  AS   A   MEANS   OF
;       IDENTIFYING   THE  SEQUENCE  NUMBER  OF  THE
;       RANDOMIZED OUTPUT LIST.  THE LABEL IS  GIVEN
;       BY THE WORD 'ITEM' FOLLOWED BY A NUMBER FROM
;       ONE TO SEVEN.  THE 'REFERENCE' OPTION ALLOWS
;       THE USER TO OUTPUT A REFERENCE FILE AHEAD OF
;       THE RANDOMLY OUTPUT FILE.  THE PROGRAM INPUT
;       IS   FREE-FORMAT   AND   CONTAINS  EXTENSIVE
;       ERROR-CORRECTION MATERIAL.  TWO SAMPLE RATES
;       (10KC  AND  20KC) ARE AVAILABLE.  THE OUTPUT
;       LIST, AS COMPUTED BY THE PROGRAM, IS PRINTED
;       FOR  USER REFERENCE.  UP TO SEVEN RANDOMIZED
;       FILES ARE ALLOWED (EXCLUDING  THE  REFERENCE
;       FILE).
```

```
FILE  [illegible].MAC        19-JUL-78                 PAGE  [illegible]


.MCALL  [illegible],.REGDEF,[illegible],.EXIT,.PRINT,.TTYOUT
.MCALL  [illegible],[illegible],[illegible],[illegible],[illegible],.CLOSE

        [illegible]
        .REGDEF
        [illegible]

.SBTTL  DATA AREA
[illegible]  [illegible]
BUFFER: .BLKW    [illegible]

BLKSIZ=40
MAXFLS=7                                  ;MAXIMUM NUMBER OF FILES USED.
NCHAN=17.                                 ;NUMBER OF CHANNELS AVAILABLE.

.PSECT  DATA
IO:     .BLKW   5*NCHAN                   ;NEW CHANNEL AREA
AREA:   .BLKW   10.                       ;MACRO CALL AREA
TABLE:  .BLKW   10.                       ;RANDOM ORDER TABLE
NUMFLS: .WORD   0,0                       ;NUMBER OF FILES NEEDED.
REFFLG: .WORD   0                         ;REFERENCE FLAG
[illegible]: .WORD   0                    ;[illegible] FLAG
HEADFLG:.WORD   0                         ;HEADING FLAG
[illegible]: .WORD   [illegible]
[illegible]: .WORD   0,0
[illegible]: .WORD   0,0
[illegible]: .WORD   0,0
[illegible]: .WORD   0,0
[illegible]: .WORD   0,0
TEMP:   .WORD   0,0,0,0
SELECT: .WORD   0,0
[illegible]: .RAD50  [illegible]
.NLIST  BIN
YESSTR: .ASCIZ  /YES/
NOSTR:  .ASCIZ  /NO/
        .ASCII  / /
INSTR:  .BLKB   82.
MES0:   .ASCII  /ENTER SAMPLE RATE REQUIRED:
        .BYTE   15,12
        .ASCII  /       1: 33KC SAMPLE RATE
        .BYTE   15,12
        .ASCII  /       2: 20KC SAMPLE RATE
        .BYTE   15,12
        .ASCII          3: 10KC SAMPLE RATE
        .BYTE   15,12
        .ASCII  /?/
        .BYTE   200
MES1:   .ASCII  /DO YOU WANT TO USE A REFERENCE FILE? /
        .BYTE   200
MES2:   .ASCII  /ENTER REFERENCE FILE NAME :/
```

```
        .BYTE   200
MESS3:  .ASCII  /ENTER THE NUMBER OF FILES (INCLUDING REFERENCE
        )/
        .ASCII  / YOU WISH TO USE: /
        .BYTE   200
MESS4:  .ASCII  /DO YOU WANT YOUR OUTPUT FILE LABELED? /
        .BYTE   200
MESS5:  .ASCII  /ENTER NAME OF FILE 1/
        .BYTE   200
MESS11: .ASCII  /ENTER INTRA-STIMULUS INTERVAL (SECONDS) : /
        .BYTE   200
MESS6:  .ASCII  /ENTER INTER-STIMULUS INTERVAL (SECONDS) : /
        .BYTE   200
MESS12: .ASCII  /DO YOU WANT TO OUTPUT A TAPE HEADING? /
        .BYTE   200
MESS13: .ASCII  /ENTER HEADING FILE NAME  */
        .BYTE   200
MESS14: .ASCII  /TYPE CARRIAGE RETURN TO PROCEED. /
        .BYTE   200
MESS7:  .ASCII  /OUTPUT ORDER: /
        .BYTE   200
MESS8:  .ASCII  /, TYPE CARRIAGE RETURN TO PROCEED.
        .BYTE   200
MESS9:  .ASCII  /DO YOU WANT ANOTHER OUTPUT SET? /
        .BYTE   200
MESS10: .ASCIZ  /DONE./
ERR1:   .ASCII  /FILE NAME ERROR; RE/
        .BYTE   200
ERR2:   .ASCII  /PLEASE ANSWER "YES" OR "NO".  /
        .BYTE   200
ERR3:   .ASCIZ  /ERROR: NUMBER OF FILES MUST BE > 2 AND < 1/
ERR4:   .ASCIZ  /FATAL ERROR: LABEL FILES NOT ON SYSTEM DISK/
ERR5:   .ASCII  /INTERVAL MUST BE GREATER THAN OR EQUAL TO/
        .ASCIZ  / 0.05 SECONDS./
ERR6:   .ASCII  /ERROR, RE/
        .BYTE   200
ERR7:   .ASCIZ  /FATAL ERROR: DAC TIME-OUT./
ITEM:   .ASCIZ  /ITEM END/
ONE:    .ASCIZ  /ONE.END/
TWO:    .ASCIZ  /TWO.END/
THREE:  .ASCIZ  /THREE.END/
FOUR:   .ASCIZ  /FOUR.END/
FIVE:   .ASCIZ  /FIVE.END/
SIX:    .ASCIZ  /SIX.END/
SEVEN:  .ASCIZ  /SEVEN.END/
.LIST   SEQ
.EVEN
```

```
FILE  RPLAY.MAC      [illegible]                     PAGE   3


.SBTTL  MAINLINE PROGRAM
.ENABL? [illegible]
BEGIN:
        MOV     #BUFFER,R4?
        .SEG?                        ;[illegible]
        .RANMIZ?                     ;[illegible] RANDOM [illegible]
        ;
        .IOPEN  [illegible]
;       SELECT THE SAMPLE FILE
        LOOP
        .PRINT  #MES0
        .INPUT  #INSTR
        .STRING #AREA,#INSTR,SELECT
        CLR     R0                   ;R0 IS ERROR FLAG
         IF <SELECT GT #3. OR SELECT LT #[illegible]>
         MOV    #-1,R0
         END
         IF <SELECT+2 NE #0>
         MOV    #-1,R0
         END
        DONEIF  <R0 EQ #0>
        .PRINT  #ERR0
        END


;       IS A REFERENCE FILE WANTED?
        LOOP
        .PRINT  #MES1
        .INPUT  #INSTR
        .FINDS  #AREA,#YESSTR,#INSTR[illegible]
         IF <R0 NE #0>
         MOV    #-1,REFFLG
         END
        DONEIF  R0 NE #0
        .FINDS  #AREA,#NOSTR,#INSTR[illegible]
         IF <R0 NE #0>
         CLR    REFFLG
         END
        DONEIF  R0 NE #0
        .PRINT  #ERR1
        END

;       IF REFERENCE IS REQUIRED, OPEN IT ON CHANNEL #0
        IF <REFFLG NE #0>
         LOOP
         .PRINT #MES2
         .INPUT #INSTR
         .OPEN  #AREA,#0,#DEFEXT,#INSTR
         ROR    R0                   ;CARRY SET INDICATES ERROR
```

```
         DONEIF <R0 GE #0>
         .PRINT #ERR1                  ;PRINT ERROR MESSAGE
         END
        END

;       GET NUMBER OF FILES, AND OPEN THEM ON CHANNELS
;       ONE THROUGH SEVEN.
        LOOP
        .PRINT  #MES3
        .INPUT  #INSTR
        .STRINT #AREA,#INSTR,#NUMFLS
        CLR     R0                     ;R0 IS ERROR FLAG
         IF <<NUMFLS LT #2> OR <NUMFLS GT #7>>
         MOV    #-1,R0
         END
         IF <NUMFLS+2 NE #0>
         MOV    #-1,R0
         END
        DONEIF  <R0 EQ #0>
        .PRINT  #ERR3
        END

;       OPEN NUMFLS FILES ON CHANNELS ONE THROUGH NUMFLS
        MOV     NUMFLS,R1                       ;R1 IS THE COUNTER
        MOV     #1,R2                           ;R2 IS THE CHANNEL NUMB
        ER
        LOOP    R1
         LOOP
         .PRINT #MES5                           ;PRINT MESSAGE
         MOV    R2,R0                           ;GET CHANNEL NUMBER
         ADD    #60,R0                          ;CONVERT TO ASCII
         .TTYOUT                                        ;PRINT IT
         .TTYOUT          #BLANK
         .TTYOUT          #'*
         .INPUT #INSTR
         .OPEN  #AREA,R2,#DEFEXT,#INSTR
         ROR    R0
         DONEIF <R0 GE #0>
         .PRINT #ERR1                           ;ERROR MESSAGE
         END
        INC     R2                              ;NEXT CHANNEL
        END

;       IS VOICED LABELING WANTED?
        LOOP
        .PRINT  #MES4
        .INPUT  #INSTR
        .FINDS  #AREA,#YESSTR,#INSTR-1
         IF <R0 NE #0>
         MOV    #-1,VLABEL
         END
```

```
 FILE  RPLAY.MAC       19-JUL-78               PAGE    7


        DONEIF  R0 NE #0
        .FINDF  #AREA,#NOSTR,#INSTR-1
         IF (R0 NE #0)
         CLR    VLABEL
         END
        DONEIF  R0 NE #0
        .PRINT  #ERR2
        END

;       IF VOICED LABELING IS DESIRED OPEN LABELING FILES
;       ON CHANNELS 8.-15.

        IF  VLABEL EQ #0                       ;IF NOT REQ., JMP
        JMP     NOLABEL
        END
                                               ;TO NOLABEL
        CLR     R1                             ;R1 IS ERROR FLAG
        .OPEN   #AREA,#8.,#DEFEXT,#ITEM
        ROR     R2
        BIS     R2,R1                          ;SET IF ERROR
        .OPEN   #AREA,#9.,#DEFEXT,#ONE
        ROR     R2
        BIS     R2,R1
        .OPEN   #AREA,#10.,#DEFEXT,#TWO
        ROR     R2
        BIS     R2,R1
        .OPEN   #AREA,#11.,#DEFEXT,#THREE
        ROR     R2
        BIS     R2,R1
        .OPEN   #AREA,#12.,#DEFEXT,#FOUR
        ROR     R2
        BIS     R2,R1
        .OPEN   #AREA,#13.,#DEFEXT,#FIVE
        ROR     R2
        BIS     R2,R1
        .OPEN   #AREA,#14.,#DEFEXT,#SIX
        ROR     R2
        BIS     R2,R1
        .OPEN   #AREA,#15.,#DEFEXT,#SEVEN
        ROR     R2
        BIS     R2,R1

;       CHECK R1 FOR ERRORS
        IF  (R1 LT #0)
        .PRINT  #ERR1
        .EXIT
        END

NOLABEL:
```

```
;        GET THE INTRA-STIMULUS TIME AND FIGURE OUT
;        WTIME1
         IF <<VLABEL NE #0> OR <REFFLG NE #0>>
          LOOP
          .PRINT #MES11
          .INPUT #INSTR
          .STRFLT #AREA,#INSTR,#TEMP          ;CONVERT TO FLOATING PO
*        INT

          .POLISH
          MOF$MS
          TEMP
          MUF$MS
          CONST
          CNF$                                ;CONVERT TO DOUBLE INTE
*        GER
          MON$SM
          TEMP
          .UNPOL


          DONEIF <<TEMP+2 EQ #0> AND <TEMP HIS #15.>>
          .PRINT #ERRS
          END
         END


;        GET THE INTER-STIMULUS TIME AND FIGURE OUT
;        WTIME2
         LOOP
         .PRINT  #MES6
         .INPUT  #INSTR
         .STRFLT #AREA,#INSTR,#TEMP+4         ;CONVERT TO FLO
*        ATING POINT

         .POLISH
         MOF$MS
         TEMP+4
         MUF$MS
         CONST
         CNF$                                 ;CONVERT TO DOUBLE INTE
*        GER
         MON$SM
         TEMP+4
         .UNPOL


         DONEIF <<TEMP+6 EQ #0> AND <TEMP+4 HIS #15.>>
         .PRINT  #ERRS
         END
```

```
;       NOW PUT THE TIMES INTO THEIR APPROPRIATE
;       MEMORY LOCATIONS, DEPENDING ON THE OPTIONS
;       SPECIFIED.
        IF <VLABEL NE #0>
        MOV     TEMP+4,WTIME1
        MOV     TEMP+6,WTIME1+2
        MOV     TEMP,WTIME2
        MOV     TEMP+2,WTIME2+2
        MOV     TEMP,WTIME3
        MOV     TEMP+2,WTIME3+2
        ELSE
         IF <REFFLG NE #0>
         MOV    TEMP+4,WTIME2
         MOV    TEMP+6,WTIME2+2
         MOV    TEMP,WTIME3
         MOV    TEMP+2,WTIME3+2
         ELSE
         MOV    TEMP+4,WTIME3
         MOV    TEMP+6,WTIME3+2
         END
        END

;       JMP BACK TO HERE FOR ANOTHER RANDOMIZED OUTPUT
BACK:

;       IS A TAPE HEADING WANTED?
HEADING:
        LOOP
        .PRINT  #MES12
        .INPUT  #INSTR
        .FINDS  #AREA,#YESSTR,#INSTR-1
         IF <R0 NE #0>
         MOV    #-1,HEADFLG
         END
        DONEIF <R0 NE #0>
        .FINDS  #AREA,#NOSTR,#INSTR-1
         IF <R0 NE #0>
         CLR    HEADFLG
         END
        DONEIF <R0 NE #0>
        .PRINT  #ERR2
        END

;       IF TAPE HEADING IS REQUIRED, OPEN IT ON CHANNEL #15.
        IF <HEADFLG NE #0>
         LOOP
         .PRINT #MES13
         .INPUT #INSTR
         .OPEN  #AREA,#15.,,#DEFEXT,#INSTR
         ROR    R0                      ;CARRY SET INDICATES ERROR
```

```
FILE  RELAY.MAC      19-JUL-78               PAGE   10

         .ENDC
          .CONDIF (R0 GE #3)
          .PRINT #ERR1             ;PRINT ERROR MESSAGE
          END
          .PRINT #MESS14
          .INPUT
          .READW #BUFFER,#15.,RELECT
          .CLOSE #15.
          JMP    HEADING
         END


;        THIS SECTION GENERATES LIST OF NUMBER IN
;        RANDOM ORDER THE SIZE OF THE LIST IN NUMFLS
;        AND THE RESULT IS IN TABLE.

;        FILL TABLE WITH SEQUENCE FROM 1 TO NUMFLS
         MOV     #TABLE,R0                 ;RANDOM ORDER TABLE
         MOV     #1,R1                     ;LOWEST CHANNEL = 1
         MOV     NUMFLS,R2                 ;COUNTER IN R2
         LOOP    R2
         MOV     R1,(R0)+                  ;FILE TABLE
         INC     R1
         END


;        NOW RANDOMIZE THE ORDER BY EXCHANGING ITEM
;        IN THE TABLE
         MOV     #TABLE,R2                 ;RANDOMIZE TABLE ORDER
         MOV     NUMFLS,R3                 ;NUMBER OF FILES IN R4
         MOV     R3,R4                     ;LOOP COUNT IN R3
         DEC     R3                        ;NUMFLS-1 REPETITIONS
         LOOP    R3
         .RANDOM                           ;RANDOM # IN R0

                                           ;(-32768 TO 32767)
         CLC                               ;(0 TO 16384)
         ROR     R0
         MUL     R4,R0                     ;MUL BY # OF ENTRIES
         ASL     R1                        ;MUL R0:R1 BY 2
         ROL     R0
         ASL     R0                        ;MAKE R0 WORD OFFSET
         ADD     R2,R0                     ;ADD CURRENT TABLE ADDR
;
         MOV     (R0),R5                   ;EXCHANGE (R0)
         ADD  (R2)+
         MOV     (R2),(R0)
         MOV     R5,(R2)+
         DEC     R4                        ;DEC # OF ENTRIES
         END
```

```
FILE  RPLAY.MAC      19-JUL-79               PAGE   11


;        NOW PRINT THE ORDER OF EVENTS
         .PRINT  #MES7
         MOV      NUMFLE,R1                    ;COUNTER IN R1
         MOV      #TABLE,R2                    ;TABLE ADDRESS IN R2
         LOOP     R1
         MOV      (R2)+,R0                     ;GET FILE # (CHAN. #)
         ADD      #60,R0                       ;CHANGE TO ASCII
         .TTYOUT
         .TTYOUT #BLANK
         END

;        WAIT FOR CARRIAGE RETURN TO PROCEED
         .PRINT  #MES8
         .INPUT


;        NOW OUTPUT THE APPROPRIATE FILES
;        LOOP UNTIL DONE

         MOV      #TABLE,R2                    ;TABLE ADDRESS IN R3
         MOV      #1,R1                        ;FILE COUNTER


AGAIN:
          IF (VLABEL NE #0)
          .GTIM  #AREA,#OLDTIM
          .RNDAC #BUFFER,#1,SELECT,#TIMER1
          ROR      R0
           IF (R0 LT #0)
           JMP     TIMOUT
           END
          MOV      (R2)+,R3                    ;CHANNEL IN R4
          ADD      #2,R3
          .RNDAC #BUFFER,R3,SELECT
          ROR      R0
           IF (R0 LT #0)                       ;TIMEOUT ERROR
           JMP     TIMOUT
           END
          END



          IF  (REFFLG NE #0)
          .GTIM   #AREA,#OLDTIM
          .RNDAC #BUFFER,#0,SELECT,#TIMER2
          ROR      R0
           IF  (R0 LT #0)                      ;TIMEOUT ERROR
           JMP     TIMOUT
           END
          END
```

```
        .GTIM    #AREA,#OLDTIM
        .READC   #BUFFER,(R2)+,SELECT,#TIMER3
        ROR      R0
        IF (R0 LT #0)                              ;TIMEOUT ERROR
        JMP      TIMOUT
        END

;       LOOP TO AGAIN UNTIL R1=NUMFLS
        IF (R1 LT NUMFLS)
        INC      R1
        JMP      AGAIN
        END


;       ASK ME NO QUESTIONS, I'LL TELL YOU NO LIES!
;       IS ANOTHER OUTPUT SET WANTED?
        LOOP
        .PRINT   #MES9
        .INPUT   #INSTR
        .FINDS   #AREA,#YESSTR,#INSTR-1
         IF (R0 NE #0)
         JMP     BACK
         END
        .FINDS   #AREA,#NOSTR,#INSTR-1
         IF (R0 NE #0)
         .SRESET
         CLR     R0
         .PRINT #MES10
         .EXIT
         END
        .PRINT   #ERR1
        END
```

```
  FILE  SPLAY.MAC      19-JUL-79              PAGE   13


 .SBTTL  SUBROUTINE AREA
TIMOUT: .PRINT  #ERAT
        .EXIT

;       TIMING ROUTINES
TIMER1:
        MOV     OLDTIM,NEWTIM                  ;HIGH PART
        MOV     OLDTIM+2,NEWTIM+2              ;LOW PART
        ADD     WTIME1,NEWTIM+2
        ADC     NEWTIM
        ADD     WTIME1+2,NEWTIM
        JMP     CHECK


TIMER2:
        MOV     OLDTIM,NEWTIM                  ;HIGH PART
        MOV     OLDTIM+2,NEWTIM+2              ;LOW PART
        ADD     WTIME2,NEWTIM+2
        ADC     NEWTIM
        ADD     WTIME2+2,NEWTIM
        JMP     CHECK

TIMER3:
        MOV     OLDTIM,NEWTIM                  ;HIGH PART
        MOV     OLDTIM+2,NEWTIM+2              ;LOW PART
        ADD     WTIME3,NEWTIM+2
        ADC     NEWTIM
        ADD     WTIME3+2,NEWTIM


CHECK:
;       NOW THE TIME-OUT TIME IS IN NEWTIM
;       COMPENSATE FOR MIDNIGHT
        IF NEWTIM HI #79.
        SUB     #6656.,NEWTIM+2
        SBC     NEWTIM
        SUB     #79.,NEWTIM
        END
        IF <NEWTIM EQ #79.> AND <NEWTIM+2 HI #6656.>
        SUB     #6656.,NEWTIM+2
        SBC     NEWTIM
        SUB     #79.,NEWTIM
        END

;       WAIT FOR CURRENT TIME TO EQUAL NEWTIM
;       CURRENT TIME IN TEMP
        LOOP
        .GTIM   #AREA,#TEMP
        IF <TEMP EQ NEWTIM> AND <TEMP+2 EQ NEWTIM+2>
        RTS     PC
```

```
FILE [illegible].MAC        [illegible]-JUL-[illegible]            PAGE  14

        END
        END

        LINES    [illegible]
```

APPENDIX F

GSCALE: COMPUTER PROGRAM USED TO MAKE ADJUSTMENTS IN PLAYBACK LEVEL

```
 FILE   GSCALE.MAC      19-JUL-79               PAGE    1


.TITLE  GSCALE

.MCALL  ..V2..,.REGDEF,.READW,.WRITW,.STRUCT
.MCALL  .PRINT,.EXIT

        ..V2..
        .REGDEF
        .STRUCT




;
;       R0:1 IS ARGUMENT
;
;       R2:3 IS SCRATCH



.MACRO  SHIFT0
        ASR     R0
        ROR     R1
.ENDM   SHIFT0

.MACRO  SHIFT2
        ASR     R2
        ROR     R3
.ENDM   SHIFT2

.MACRO  DADD
        ADD     R3,R1
        ADC     R0
        ADD     R2,R0
.ENDM   DADD



DBO:    RTS     PC
DBO1:   SHIFT0
        MOV     R0,R2
        MOV     R1,R3
        SHIFT2
        DADD
        SHIFT2
        DADD
        SHIFT2
        SHIFT2
        SHIFT2
```

```
         DADD
         RTS     PC
D302:    SHIFT0
         MOV     R0,R2
         MOV     R1,R3
         SHIFT2
         DADD
         SHIFT2
         SHIFT2
         SHIFT2
         DADD
         SHIFT2
         DADD
         RTS     PC
D303:    SHIFT0
         MOV     R0,R2
         MOV     R1,R3
         SHIFT2
         SHIFT2
         DADD
         SHIFT2
         DADD
         SHIFT2
         SHIFT2
         DADD
         RTS     PC
D304:    SHIFT0
         MOV     R0,R2
         MOV     R1,R3
         SHIFT2
         SHIFT2
         DADD
         RTS     PC
D305:    SHIFT0
         MOV     R0,R2
         MOV     R1,R3
         SHIFT2
         SHIFT2
         SHIFT2
         DADD
         RTS     PC
D306:    SHIFT0
.        RTS     PC
D307:    SHIFT0
         JSR     PC,D601
         RTS     PC
D308:    SHIFT0
         JSR     PC,D602
         RTS     PC
D309:    SHIFT0
         JSR     PC,D603
         RTS     PC
```

```
0BU0:   SHIFT0
        JSR     PC,0BD4
        RTS     PC
0BU1:   JSR     PC,0BD5
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU2:   JSR     PC,0BD4
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU3:   JSR     PC,0BD3
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU4:   JSR     PC,0BD2
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU5:   JSR     PC,0BD1
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU6:   ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU7:   JSR     PC,0BD5
        ASL     R1
        ROL     R0
        BVS     OVER
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU8:   JSR     PC,0BD4
        ASL     R1
        ROL     R0
        BVS     OVER
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
0BU9:   JSR     PC,0BD3
        ASL     R1
        ROL     R0
        BVS     OVER
```

```
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
DBU10:  JSR     PC,DBU2
        ASL     R1
        ROL     R0
        BVS     OVER
        ASL     R1
        ROL     R0
        BVS     OVER
        RTS     PC
OVER:   BCS     1$
        MOV     #77777,R0
        MOV     #177777,R1
        RTS     PC
1$:     MOV     #177777,R0
        MOV     #177777,R1
        RTS     PC
;
;       JUMP TABLE
;
        DBD10
        DBD9
        DBD8
        DBD7
        DBD6
        DBD5
        DBD4
        DBD3
        DBD2
        DBD1
JTABLE: DB0
        DBU1
        DBU2
        DBU3
        DBU4
        DBU5
        DBU6
        DBU7
        DBU8
        DBU9
        DBU10


YESNO:
        LOOP
        .INPUT  #INBUF
        .FINDS  #AREA,#YES,#INBUF-1
        MOV     #-1,YESFLG
        DONEIF  R0 NE #0
        .FINDS  #AREA,#NO,#INBUF-1
```

```
        CLR     YESFLG
        DONEIF  R0 NE #0>
        .PRINT  #ERR1
        END
        RTS     PC


MYROT1: .PRINT  #MESS3
        .INPUT
        RTS     PC
MYROT2: .PRINT  #MESS5
        .INPUT
        RTS     PC

.NLIST  BIN
ERR1:   .ASCII  /PLEASE TYPE YES OR NO? /
        .BYTE   200
ERR2:   .ASCII  /FILE NAME ERROR.  RE/
        .BYTE   200
ERR3:   .ASCII  /ERROR: MUST BE >= -10 AND <=10  /
        .BYTE   15,12
        .ASCII  /RE/
        .BYTE   200
ERR4:   .ASCIZ  /FATAL ERROR: SCRATCH FILE NOT LARGE ENOUGH./
MESS1:  .ASCII  /ENTER FILE NAME /
        .BYTE   200
MESS2:  .ASCII  /PLAY AT ORIGINAL LEVEL? /
        .BYTE   200
MESS3:  .ASCII  /TYPE <CR> TO PLAY AT ORIGINAL LEVEL./
        .BYTE   200
MESS4:  .ASCII  /ENTER SCALE FACTOR IN DB? /
        .BYTE   200
MESS5:  .ASCII  /TYPE <CR> TO PLAY AT NEW LEVEL./
        .BYTE   200
MESS6.  .ASCII  /LEVEL O.K.? /
        .BYTE   200
JUNKFL: .ASCIZ  /JUNK.JNK/
        .ASCII  / /
INBUF:  .BLKB   81.
YES:    .ASCIZ  / YES/
NO:     .ASCIZ  / NO/
.LIST   BIN
.EVEN
VAR:    .WORD   0,0
AREA:   .BLKW   12.
YESFLG: .WORD   0
FSIZE:  .WORD   0
NUMBLK: .WORD   0
NXTBLK: .WORD   0
OUTHC:  .WORD   0
RSIZE:  .WORD   0
DEFEXT: .RAD50  /SND/
```

```
BUFFER: .BLKW   6*1024.

BEGIN:
        LOOP
        .PRINT  #MES1
        .OPEN   #AREA,#0,#IEFEXT
        ROR     R1
        MOV     R0,FSIZE
        DONEIF <R1 GE #0>
        .PRINT  #ERR2
        END

        LOOP
        .PRINT  #MES2
        JSR     PC,YESNO
        DONEIF <YESFLG EQ #0>
        .RNDAC  #BUFFER,#0,#3,#M,ROT1
        END


        .OPEN   #AREA,#1,,#JUNKFL
        ROR     R1
        IF <<R0 LT FSIZE> OR <R1 LT #0>>
        .PRINT  #ERR4
        .EXIT
        END

BACK:
        LOOP
        .PRINT  #MES4
        .INPUT  #INBUF
        .STRINT #AREA,#INBUF,#VAR
        DONEIF <<VAR LE #10.> AND <VAR GE #-10.>>
        .PRINT  #ERR3
        END

;       SELECT ROUTINE
        MOV     VAR,R0
        ASL     R0
        MOV     JTABLE(R0),ROUTAD
        MOV     FSIZE,NUMBLK
        CLR     NXTBLK

        LOOP
        DONEIF <NUMBLK LE #0>
        IF <NUMBLK GE #24.>
        MOV     #24.,R4
        ELSE
        MOV     NUMBLK,R4
        END
        MOV     R4,R5
        SWAB    R5
```

```
        .READW  #AREA,#0,#BUFFER,R5,NXTBLK
        MOV     R5,BSIZE
        MOV     #BUFFER,R4
        LOOP    R5
        MOV     (R4),R0
        CLR     R1
        JSR     PC,@ROUTAD
        MOV     R0,(R4)+
        END
        MOV     BSIZE,R5
        MOV     R5,R4
        SWAB    R4
        .WRITW  #AREA,#1,#BUFFER,R5,NXTBLK
        ADD     R4,NXTBLK
        SUB     R4,NUMBLK
        END

        MOV     FSIZE,R1
        DEC     R1
        .RNDAC  #BUFFER,#1,#3,#MYROT2,#0,R1


        .PRINT  #MESS
        JSR     PC,YESNO
        IF <YESFLG EQ #0>
        JMP     BACK
        END

        MOV     FSIZE,NUMBLK
        CLR     NXTBLK

        LOOP
        DONEIF <NUMBLK LE #0>
        IF <NUMBLK GE #24.>
        MOV     #24.,R4
        ELSE
        MOV     NUMBLK,R4
        END
        MOV     R4,R5
        SWAB    R5
        .READW  #AREA,#1,#BUFFER,R5,NXTBLK
        .WRITW  #AREA,#0,#BUFFER,R5,NXTBLK
        ADD     R4,NXTBLK
        SUB     R4,NUMBLK
        END

        .EXIT


        .END    BEGIN
```

APPENDIX G

FREQUENCY RESPONSE MEASUREMENTS ON EQUIPMENT USED TO PREPARE AND PRESENT SPEECH STIMULI

The frequency responses of filtered speech stimuli are often subject not only to intentional filtering, but to unintentional filtering as well. Unintentional filtering may be imposed by all instruments used in stimulus preparation and presentation. The purpose here is to describe the apparatus and procedures used to obtain response curves on (1) the Ampex AG-500 recorder, (2) the Nakamichi 700 II recorder, (3) a computer system, and (4) the Grason-Stadler 162 speech audiometer with TDH-49 earphones mounted with standard cushions (MX 41/AR). For measurement purposes, the upper frequency limit was arbitrarily set at 10 kHz for all systems.

Tape Recorders

Figure G-1 describes the apparatus used to measure the frequency response of each of the two tape recorders. The input to the tape recorders was produced by a Wavetek 185 sine generator and was monitored by a Ballantine 5500 B frequency counter. The output of the recorders was monitored by a Bruel and Kjaer 2607 measuring amplifier. Initially, the sine generator was adjusted to emit a 1 kHz signal and the Nakamichi recorder was adjusted to zero VU and 580 mv output. Subsequently, readings were made from the measuring amplifier in dB relative to an arbitrary reference. A similar procedure was used with the Ampex recorder, but the Ampex recorder had a 1.5 v output when adjusted for zero VU. The response curves obtained in this manner are shown in Figure G-2.

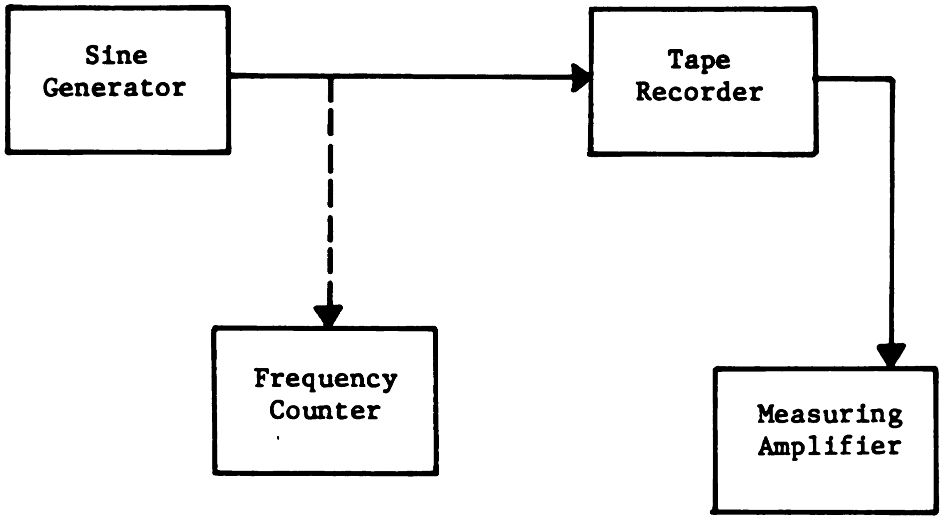

Figure G-1. Apparatus used for measuring the frequency response of the tape recorders.

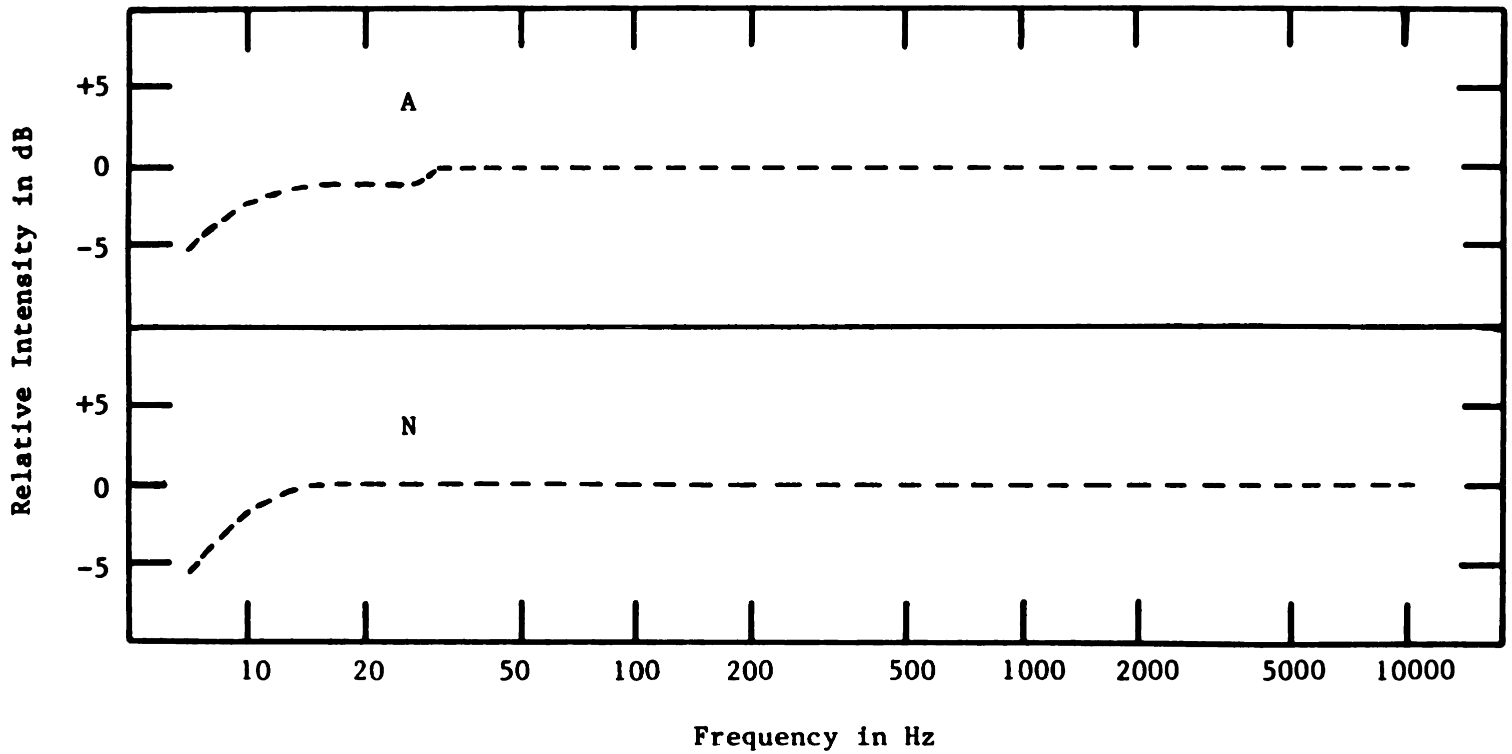

Figure G-2. Lower portion of the frequency responses for the Ampex AG-500 recorder (A) and the Nakamichi 700 II recorder (N).

Computer System

Figure G-3 describes the apparatus used for measuring the frequency response of the computer system. The Hewlett-Packard 3310 A sine generator provided the input signal to a 3 Rivers Computer Corporation analog-to-digital converter (ADC). The input signal was monitored by a frequency counter (Heath/Schlumberger SM 4100) and an electronic voltmeter (Bruel and Kjaer 2409). The output of the ADC was passed to a Digital Equipment Corporation pdp 11/40 computer which was interfaced with two RK05 disk drives, a decwriter II input-output terminal, and a video monitor. The output of the computer was sent to a 3 Rivers Computer Corporation digital-to-analog converter (DAC). Finally, the DAC output was monitored by a measuring amplifier (Bruel and Kjaer 2607) where the intensity was read in decibels (relative to an arbitrary reference) as a function of sinusoidal inputs. Prior to taking measurements, the sine generator was adjusted to emit a 250 mv amplitude for a 1 kHz signal, and the ADC was set for a sampling rate of 10 kHz. The response curve obtained on the computer system is shown in Figure G-4.

Speech Audiometer and Earphones

Figure G-5 shows the apparatus used to measure the combined frequency response of the Grason-Stadler speech audiometer and the TDH-49 earphones. A swept sine generator (Bruel and Kjaer 1024) supplied the input to the speech audiometer. The input was monitored by a frequency counter (Ballantine 5500 B) and a measuring amplifier (Bruel and Kjaer 2607). The output of the audiometer was routed to one of the TDH-49 earphones. The earphone was mounted on an artificial ear (Bruel and Kjaer 4152) which contained a coupler (National Bureau of Standards 9A) and a condenser microphone (Bruel and Kjaer 4144). Finally, the output of the artificial ear was

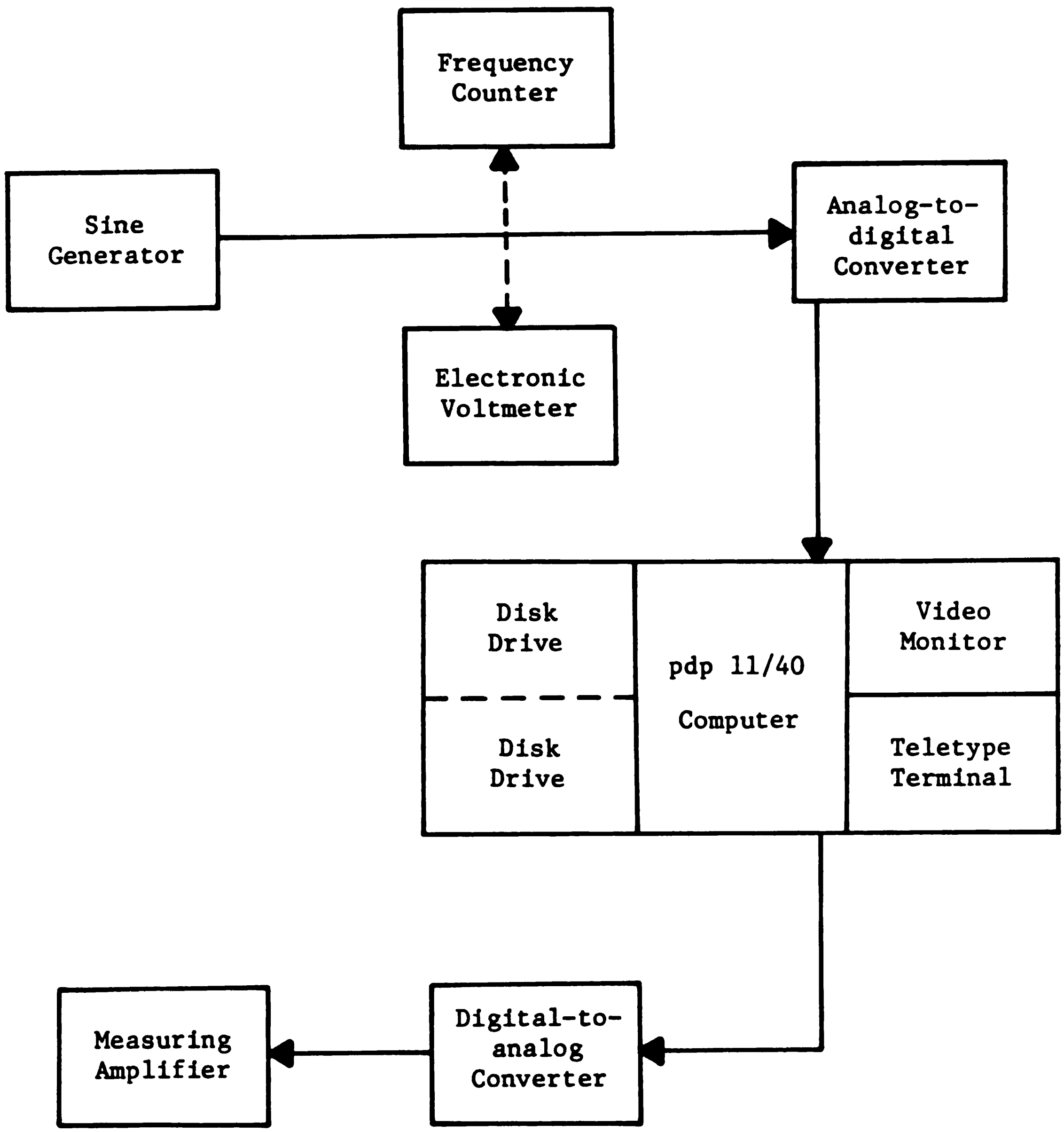

Figure G-3. Apparatus used for measuring the frequency response of the computer system.

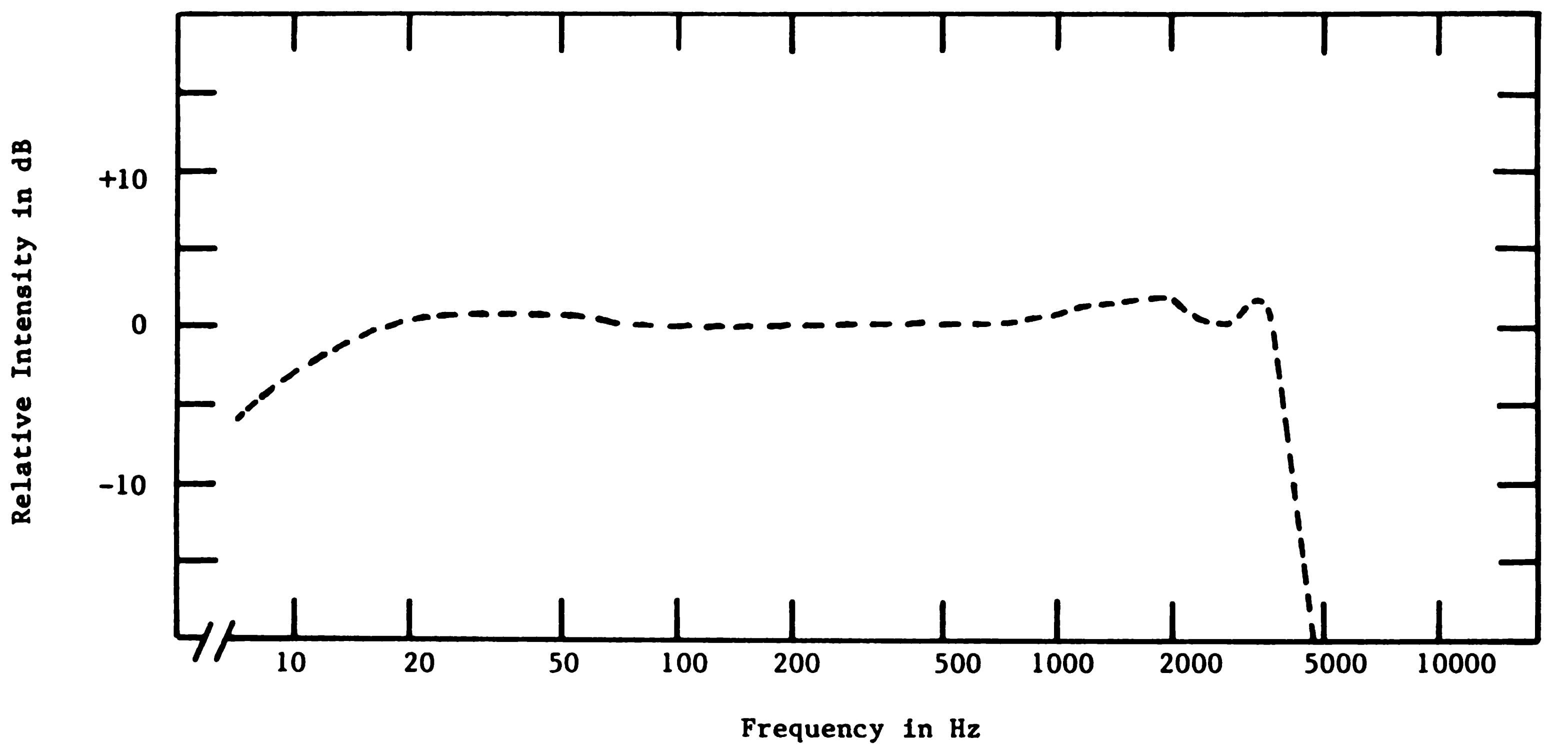

Figure G-4. Frequency response for the computer system.

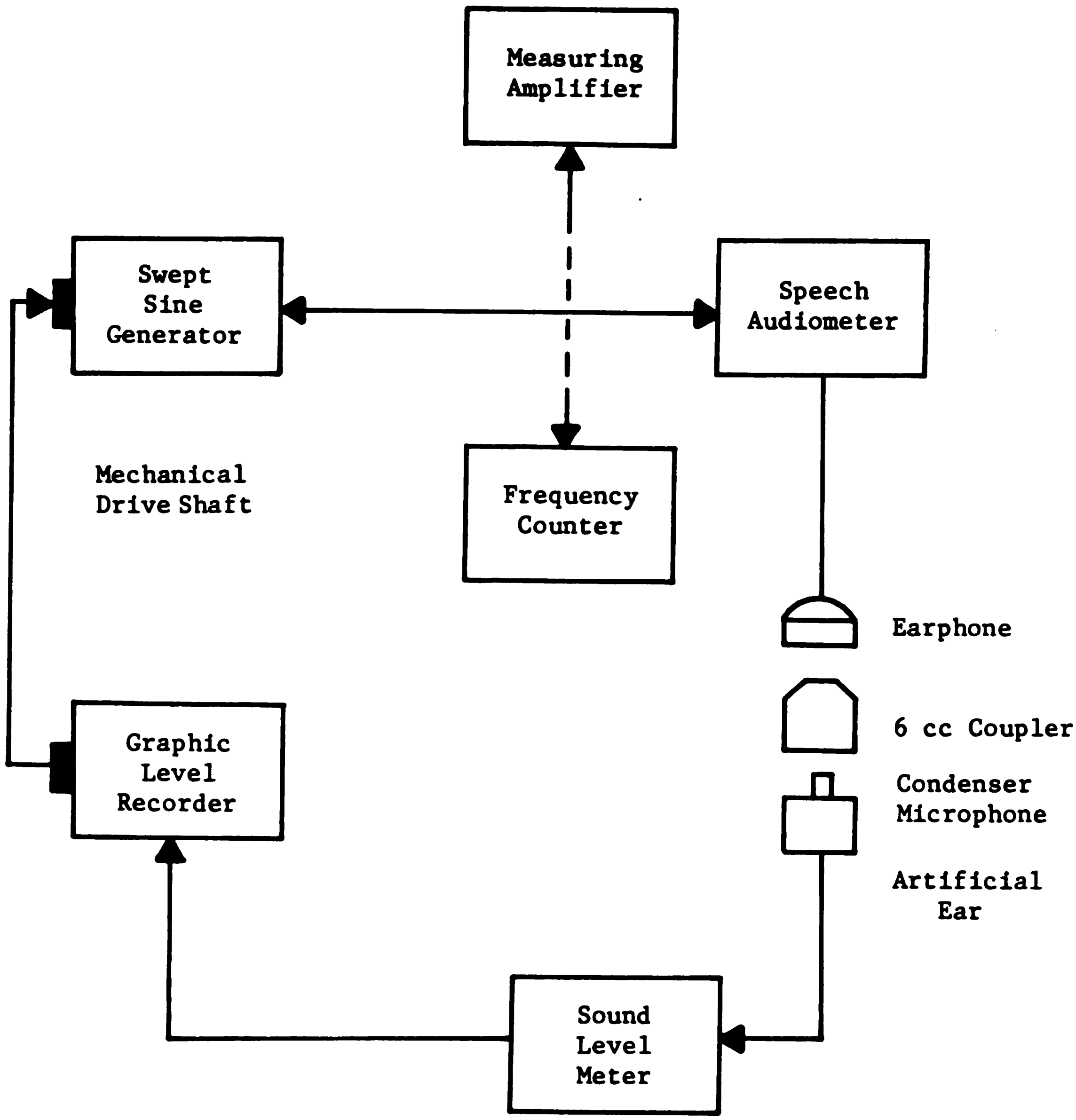

Figure G-5. Apparatus used for measuring the combined frequency response of the speech audiometer and earphones.

passed through a precision sound level meter (Bruel and Kjaer 2204) to a graphic level recorder (Bruel and Kjaer 2305). The graphic level recorder provided a synchronized mechanical drive to the sine generator.

The two earphones (E-1 and E-2) and the two channels of the audiometer were tested separately. E-1 was tested with channel 1 and E-2 was tested with channel 2. Initially, the sine generator was adjusted to emit a 1 kHz signal of 250 mv. The attenuator for the test channel was set to 80 dB HL, and the attenuator for the nontest channel was set to -10 dB HL. The signal in the test channel was adjusted to zero VU. Then, the response curve was run. For very low frequency measurements ($<$20 Hz) a Wavetek 185 sine generator was substituted for the sweep generator and discrete readings were made on the sound level meter. The response curves obtained for the speech audiometer with each of the two earphones is shown in Figure G-6.

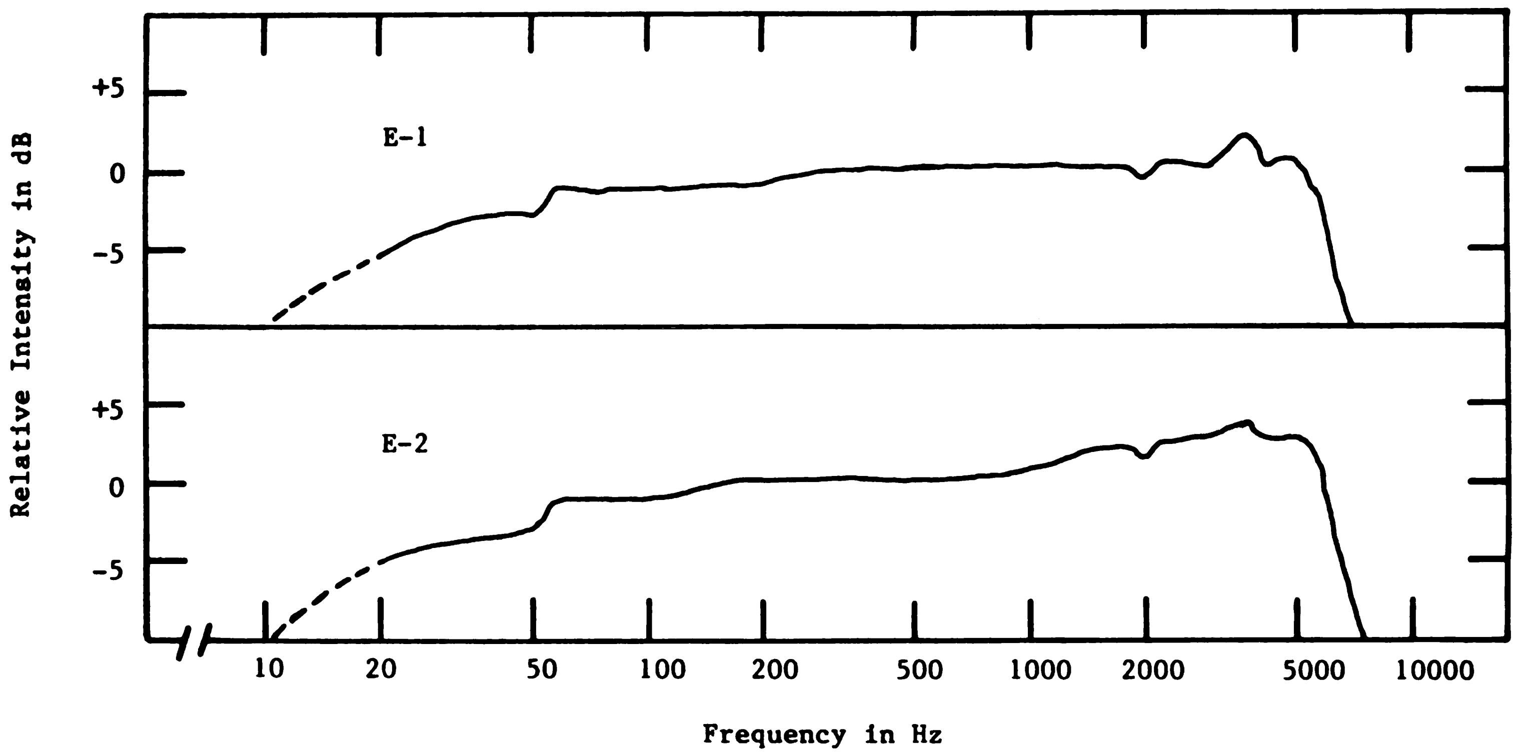

Figure G-6. Frequency response for the Grason-Stadler 162 speech audiometer and the TDH-49 earphones (E-1 and E-2).

APPENDIX H

RUN PROTOCOL

Project Name: ______________________________

Experimenter: ______________________________

<u>Subject Identification</u>

Name: ________________________________ Number: __________

Informed Consent Release Form signed? Yes__ No __

<u>Audiologic Screening Results</u>

History clear? Yes__ No __

Criteria met?

Right ear: Yes__ No __

Left ear: Yes__ No __

Experimental group assignment:

Normal listener group _____ Impaired listener group _____

<u>Visual Training and Screening</u>

Magnitude estimation training completed? Yes__ No __

Test-Retest reliability correlation (r) = _____

Same magnitude estimates always given to standard and identical comparison stimuli? Yes__ No __

Test-Retest agreement obtained for 6 of 7 pairs? Yes__ No __

Eligible to continue in the study? Yes__ No __

<u>Experimental Session #1</u>

Date: ____________ Time _________ Cal. done? Yes__ No __

Instructions given? ______________ Practice completed? __________

Experimental condition order (Tape for <u>S</u> ____________)

_____ _____ _____ _____ _____ _____
1st 2nd 3rd 4th 5th 6th

Test ear: R _____ L _____ Presentation level: _____dB HL

Subject selected for second experimental session? Yes__ No __

<u>Experimental Session #2</u>

Date: ______________ Time ________ Cal. done? Yes__ No __

Instructions given? ______________ Practice completed? __________

Experimental condition order (Tape for <u>S</u> ____________)

_____ _____ _____ _____ _____ _____
1st 2nd 3rd 4th 5th 6th

Test ear: R ______ L ______ Presentation level: _____dB HL

APPENDIX I

INFORMED CONSENT RELEASE FORM III

1. I, ________________________, freely and voluntarily consent to serve as a subject in a scientific study of speech perception conducted by Dr. Michael R. Chial, Mr. Gary Lawson, and other student assistants.

2. I understand that the purpose of the study is to determine the reliability and validity of a particular method of measurement of speech perception which may be of future clinical usefulness.

3. I understand that I will not be exposed to any experimental conditions which constitute a threat to my hearing, nor to my physical or psychological well-being.

4. I understand that data gathered from me for this experiment are confidential, that no information uniquely identified with me will be made available to other persons or agencies, and that any publication of the results of this study will maintain anonymity.

5. I engage in this study freely, without payment to me or from me, and without implication of personal benefit. I understand that I may cease participation in the study at any time.

6. I have had the opportunity to ask questions about the nature and purpose of the study, and I have been provided with a copy of this written informed consent form. I understand that upon completion of the study, and at my request, I can obtain additional explanation about the study.

Date: ________________________ Signed: ____________________________

PRO. NO. ______________________ SEQ. NO. ______________________

GRP. NO. ______________________

11/22/77

APPENDIX J

AUDIOLOGICAL SCREENING FORM

Project: ____________ Date: _________ Time: __________ Examiner: _________

Subject Identification

Name: _______________________________ Subject No. _____________________

Birthdate: ___________________________ Age: ___________________________

History

Recent onset of hearing loss?	Yes ____	No ____
Currently active URI?	Yes ____	No ____
Family history of genetic hearing loss?	Yes ____	No ____
Vertigo?	Yes ____	No ____
Tinnitus?	Yes ____	No ____
Otologic surgery?	Yes ____	No ____
Informed Release Form signed?	Yes ____	No ____

Test Results

*Pure Tone Air Conduction Thresholds (dB HL)

Frequency (Hz)		250	500	1000	2000	4000	6000	2-tone Aver.	Test Ear
Ear	R							R	R ______
	L							L	L ______

*Olsen-Noffsinger TDT (Impaired hearing Ss only)

Frequency (Hz)		1000	4000
Ear	R	Pass _____ Fail _____	Pass _____ Fail _____
	L	Pass _____ Fail _____	Pass _____ Fail _____

*Speech discrimination score in test ear at 40 dB SL relative to the 2-tone average threshold: ______ %.

*Tympanometry

Pressure (mm H_2O)		-200	-100	0	+100	+200
Compliance (CC) in Probe Ear	R					
	L					

MEP (mm H_2O)	Compliance (CC)
R	
L	

*Acoustic Reflexes (normal hearing Ss only)

		Threshold (dB HL)			Reflex Delay Test	
Frequency (Hz)		500	1000	2000	500	1000
Stimulus Ear	R				Pass ___ Fail ___	Pass ___ Fail ___
	L				Pass ___ Fail ___	Pass ___ Fail ___

*Equipment used (identify):

Audiometer ___________________________ Impedance meter _______________

APPENDIX K

SCRIPT FOR VISUAL TRAINING AND SCREENING TASK

Preliminary Instructions (Presented auditorily and visually)

Slide 1: blank	Please read these instructions as you listen to them. You are going to see some pictures of squares and circles. We want you to assign numbers to these shapes in a special way. Before we explain the numbers, let's look at a sample of what you'll be seeing.
Slide 2: two squares	Notice the square on the left. This shape is labeled "S" for standard. The "S" shape will always be on the left. The other shape is labeled "C" for comparison. This shape will always be on the right. However, we will be changing the size of the shape labeled "C".
Slide 3: one square	Here we have only one shape, a standard. When you see a slide like this, you are to assign a number to the shape. The number you pick should represent your impression of the <u>size</u> of the shape. You can pick any number you want, but it will be easier if you pick a whole number.
Slide off	In a moment we'll show you some more pairs of shapes. One will be a standard; the other will be a comparison. Your job will be to refer to the number you initially gave shape "S" and then pick a number for shape "C". The number for shape "C" should represent the size of shape "C" relative to the size of the standard. In other words, the <u>number</u> you pick for "C" should relate to the <u>number</u> for "S" in the same way that the <u>size</u> of "C" relates to the <u>size</u> of "S".
	Let's try an example. We show you a slide with a single shape on it. That shape is labeled "S" for standard. Following the instructions given at the time, you assign that standard some number, let's say 20. Then we show a slide with the same shape "S", plus a shape "C". If "C" looks twice as big as "S", then "C" gets the number 40. If "C" looks half as big as "S", then "C" gets the number 10. Each time you see a new slide, you give the "C"

shape a number that represents its size compared to the size of "S". Or, let's assume that you called shape "S" 5. If the comparison shape "C" looks three times bigger than "S", you'd call the "C" shape 15. If "C" looks one-fifth as big as the standard, you'd call it 1.

We're just about ready to start. You'll be seeing three sets of shapes. The first set is for practice and consists of squares. The next two sets are test items and use circles. Remember, this is not an intelligence test or a "trick" test. Even though there are no wrong answers, we want you to pay careful attention to what you see and hear. Remember, your job is to recall the number you gave "S" and to assign a number to "C" that represents the size of "C" compared to the size of "S".

If you want to reread these instructions or ask the experimenter a question, press the button marked "STOP".

When you are ready to begin, press the button marked "PLAY" and go to the next page.

Once you start, do not press the "STOP" button until you are told to do so.

Introduction to Response Sheet (Presented Auditorily)

Look at your response sheet. Enter the numbers you assign to shapes in the boxes, starting at the top of each column.

Slides for Stimulus Set "A" (Trial 1)

This is Trial 1. Use the lefthand column of your response sheet.

Slide 4: audio instructions

Please assign a number to the shape labeled "S" in the next slide. You have 10 seconds.

Slide 5: one square

(10 seconds)

Be sure your answer is on the response sheet.

Slide 6: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 7: two squares

(10 seconds)

Be sure your answer is on the response sheet.

Slide 8: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 9: two squares

(10 seconds)

Be sure your answer is on the response sheet.

Slide 10: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 11: two squares

(10 seconds)

Be sure your answer is on the response sheet.

Slide 12: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 13: two squares

(10 seconds)

Be sure your answer is on the response sheet.

Slide 14: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 15: two squares

(10 seconds)

Be sure your answer is on the response sheet.

Slide 16: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 17: two squares (10 seconds)

Be sure your answer is on the response sheet.

Slide 18: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 19: two squares (10 seconds)

Be sure your answer is on the response sheet. We have finished Trial 1.

Slides for Stimulus Set "B" (Trial 2)

Slide off

This is Trial 2. Use the middle column of your response sheet.

Slide 20: audio instructions

Please assign a number to the shape labeled "S" in the next slide. You have 10 seconds.

Slide 21: one circle (10 seconds)

Be sure your answer is on the response sheet.

Slide 22: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 23: two circles (10 seconds)

Be sure your answer is on the response sheet.

Slide 24: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 25: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 26: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 27: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 28: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 29: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 30: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 31: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 32: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 33: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 34: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 35: two circles
(10 seconds)

Be sure your answer is on the response sheet. We have finished Trial 2.

Slides for Stimulus Set "C" (Trial 3)

Slide off

This is Trial 3. Use the righthand column of your response sheet.

Slide 36: audio instructions

Please assign the <u>same number</u> you used for the <u>last shape labeled</u> "S" to the shape labeled "S" in the next slide.

Slide 37: one circle
(10 seconds)

Be sure your answer is on the response sheet.

Slide 38: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 39: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 40: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 41: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 42: audio instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 43: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 44: audio
instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 45: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 46: audio
instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 47: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 48: audio
instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 49: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide 50: audio
instructions

Please assign a number to the shape labeled "C" in the next slide. You have 10 seconds.

Slide 51: two circles
(10 seconds)

Be sure your answer is on the response sheet.

Slide off

You may now press the button marked "STOP". Inform the experimenter that you have completed this part of the experiment.

APPENDIX L

INSTRUCTIONS AND RESPONSE SHEET FOR VISUAL TASKS

Please read these instructions as you listen to them. You are going to see some pictures of squares and circles. We want you to assign numbers to these shapes in a special way. Before we explain the numbers, let's look at a sample of what you'll be seeing.

Slide of Two Squares. Notice the square on the left. This shape is labeled "S" for standard. The "S" will always be on the left. The other shape is labeled "C" for comparison. This shape will always be on the right. However, we will be changing the size of the shape labeled "C".

Slide of One Square. Here we have only one shape, a standard. When you see a slide like this, you are to assign a number to the shape. The number you pick should represent your impression of the size of the shape. You can pick any number you want, but it will be easier if you pick a whole number. (slide off)

In a moment, we'll show you some more pairs of shapes. One will be a standard; the other will be a comparison. Your job will be to refer to the number you initially gave the shape "S" and then pick a number for shape "C". The number for shape "C" should represent the size of shape "C" relative to the size of the standard. In other words, the number you pick for "C" should relate to the number for "S" in the same way that the size of "C" relates to the size of "S".

Let's try an example. We show you a slide with a single shape on it. That shape is labeled "S" for standard. Following the instructions given at the time, you assign that standard some number, let's say 20. Then we show a slide with the same shape "S", plus a shape "C". If "C" looks twice as big as "S", then "C" gets the number 40. If "C" looks half as big as "S", then "C" gets the number 10. Each time you see a new slide, you give the "C" shape a number that represents its size compared to the size of "S". Or, let's assume that you called shape "S" 5. If the comparison shape "C" looks three times bigger than "S", you'd call the "C" shape 15. If "C" looks one-fifth as big as the standard, you'd call it 1.

We're just about ready to start. You'll be seeing three sets of shapes. The first set is for practice and consists of squares. The next two sets are test items and use circles. Remember, this is not an intelligence test or a "trick" test. Even though there are no wrong answers, we want you to pay careful attention to what you see and hear. Remember, your job is to recall the number you gave "S" and to assign a number to "C" that represents the size of "C" compared to the size of "S".

If you want to reread these instructions or ask the experimenter a question, press the button marked "STOP".

When you are ready to begin, press the button marked "PLAY" and go to the next page.

Once you start, do not press the "STOP" button until you are told to do so.

Page_____ of_____

RESPONSE SHEET

Subject No.____________ Date_____________ Session______________

Stimulus Sets		
Trial 1	Trial 2	Trial 3
S	S	S
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7

APPENDIX M

PILOT STUDY OF A VISUAL MAGNITUDE ESTIMATION TASK

A 15-minute program of audio-taped instructions and stimulus slides was developed to quickly teach initially naive subjects how to perform a magnitude estimation task and to enable the examiner to check test-retest reliability.

The slides presented three sets of stimuli. Each set represented a trial comprised of seven pairs of stimuli. Each pair consisted of two white geometric forms located side by side on a blue background and was produced on a 2" X 2" slide. Each slide had an identifying number centered beneath the two stimuli. The left member of the pair, the standard stimulus, was located below an "S" for standard. The right member of the pair, the comparison stimulus, was located below a "C" for comparison. Although the size of the standard did not change within a trial, each of the seven comparisons differed in size. The middle-sized comparison was the same size as the standard. The first trial consisted of randomly ordered squares of different areas in square inches (i.e., 0.063, 0.250, 0.563, 0.766, 1.000, 1.266, and 1.891). The second and third trials consisted of different random orders of circles of different areas in square inches (i.e.,
1 2 3 4 5 6 7
0.049, 0.196, 0.442, 0.601, 0.754, 0.994, 1.485).

Purpose

This study was designed to obtain pilot data on the newly developed materials. The data were examined to answer the following questions:

1. Is there a statistically significant correlation between each subject's magnitude estimates on Trial 2 and Trial 3 after a single training trial (i.e., Trial 1)?

2. Is there a statistically significant trend for the log geometric

mean magnitude estimates across Trials 2 and 3 to be influenced by changes in log circle size (in^2)?

3. If the log geometric mean magnitude estimates across Trials 2 and 3 are significantly different as a function of log circle size, is the linear component of the trend statistically significant?

4. Is there a statistically significant correlation between the slopes of the log magnitude estimates for Trial 2 and Trial 3 when they are examined as a function of log circle size (in^2)?

Method

The visual magnitude estimation task was administered to 12 college students who reported no hearing loss or uncorrected vision problems. Each subject's job was to estimate the apparent magnitude of various squares and circles projected on a rear screen (9" by 9") slide viewer with a synchronized tape player (Singer Caramate II SP). The audio signal was presented through earphones at a comfortable loudness level. Each subject was allowed to assign any value to the standard stimulus, but was asked not to change this value within a trial. Subjects responded by marking their estimated numerical value of the comparison stimulus relative to the value of the standard stimulus on a response sheet.

The first stimulus set (squares) was used to clarify the subject's concept of magnitude estimation. The second and third stimulus sets (circles) were used to gather data for checking test-retest reliability.

Results and Analysis

The experimental questions were based on three types of dependent variables obtained for each each subject: (1) magnitude estimates for Trial 2 and Trial 3, (2) log geometric mean magnitude estimates across Trials 2 and 3, and (3) the slopes of the least squares lines of best fit for the log perceptual values as a function of log circle size (in^2). The slopes were obtained for log magnitude estimates on Trial 2 and Trial 3 as a function of log circle size (in^2) and for log geometric mean magnitude estimates across Trials 2 and 3 as a function of log circle size (in^2). The magnitude estimates were available as raw data. The other dependent variables were obtained as output from MAGEST, a Fortran IV computer program for analyzing magnitude estimation data (Kerst, 1978). MAGEST required the magnitude estimates and stimulus magnitudes as input.

Pearson product-moment correlation coefficients (Linton and Gallo, 1975, pp. 347-352) were computed to determine the relationship between the magnitude estimates obtained on Trials 2 and 3 for each of the 12 subjects. The 12 coefficients ranged from 0.93 to 1.00 and were all significant ($r = 0.754$ for $\alpha = 0.05$, df = 5). A coefficient of determination (Linton and Gallo, 1975, pp. 344-346) was computed for the subject with the lowest correlation coefficient ($r = 0.93$). The coefficient of determination ($r^2 = 0.86$) showed that if magnitude estimates on one trial are used to predict those on the other, 86% of the total variance in that subject's estimates can be accounted for. Also, 100% agreement was noted across subjects and trials between magnitude estimates assigned to the standard stimuli and those assigned to equivalent comparison stimuli.

Figure M-1 shows the group mean ($n = 12$) log geometric mean magnitude estimates across Trials 2 and 3 plotted as a function of log circle size (in^2). Table M-1 summarizes the results of a one-way analysis of variance for repeated measures (Winer, 1971, pp. 261-268). These results suggest a statistically significant difference in log geometric means for trials as a function of circle size ($\alpha < 0.0001$ for F observed = 226.993; df = 6, 66).

Table M-2 summarizes a test for linear trend (Winer, 1971, pp. 296-300)

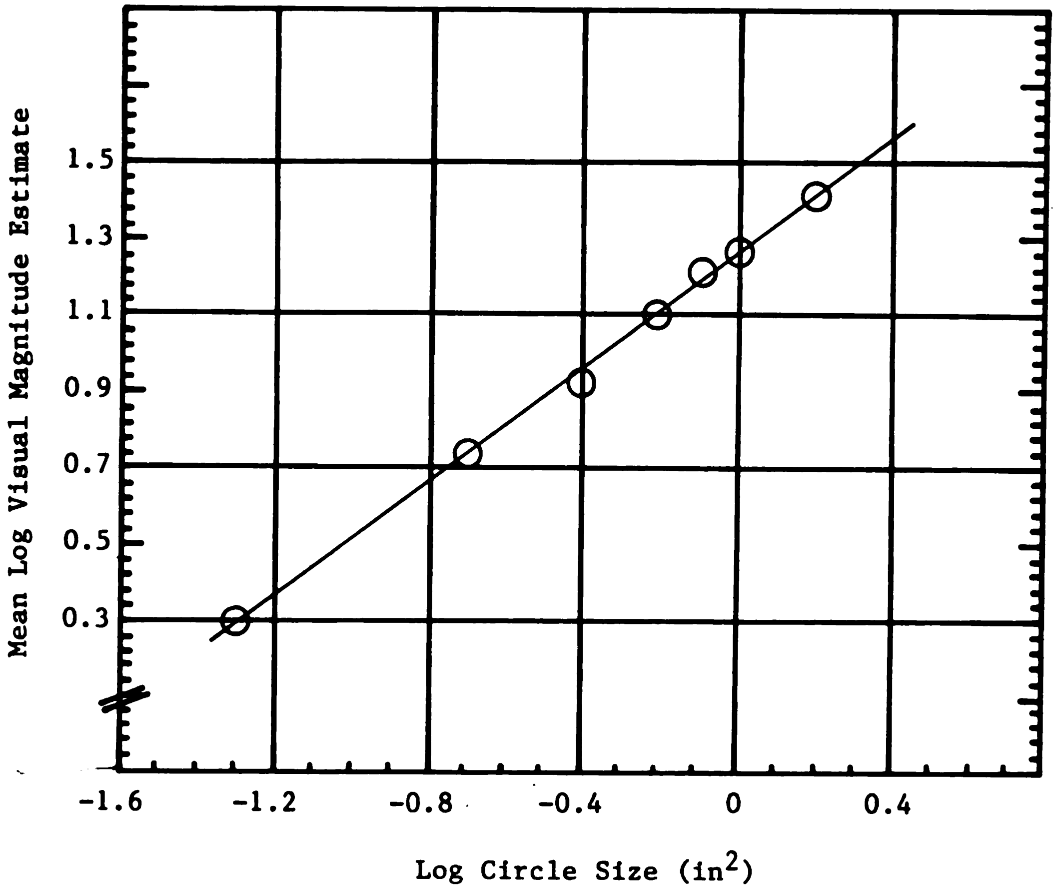

Figure M-1. Mean log geometric mean visual magnitude estimates across Trials 2 and 3 for the visual pilot group (N = 12) plotted as a function of log circle size (in^2). The solid line represents the least squares line of best fit.

Table M-1. Summary of analysis of variance in log geometric mean (G.M.) magnitude estimates (Mag. Est.) as a function of circle size.

Source	SS	df	MS	F	F for $\alpha = .05$	α for $F_{observed}$
Between subjects	1.9528	11	0.1775			
Within subjects	9.7937	72	0.1360			
G.M. Mag. Est.	9.3409	6	1.5568	226.933*	3.992*	<0.0001
Residual	0.4528	66	0.0069			
Total	11.7465	83				

*Interpolated from F distribution Table (Winer, 1971, pp. 864-869).

Table M-2. Summary of test for linear trend in log geometric mean magnitude estimates of circle size.

Source of Variance	SS	df	MS	F	F for $\alpha = .05$	α for $F_{observed}$
Linear trend	7.59	1	7.59	253*	8.5468*	< 0.0001
Deviation from linear trend	2.20	71	0.03			

*Interpolated from F distribution table (Winer, 1971, pp. 863-869).

based on the analysis of variance summarized in Table M-1. The orthogonal coefficients for this test were computed according to the method described by Kirk (1968, pp. 513-517) since Winer did not describe a procedure for unequal intervals of the independent variable. The results in Table M-2 suggest that a linear equation accounts for a statistically significant part of the trend observed in the analysis of variance ($\alpha < 0.0001$ for $F_{\text{linear trend}} = 253$; df = 1, 71). The portion of variance accounted for by the linear component is approximately 81%.

The slopes (Table M-3) for the linear functions described above (i.e., log geometric mean magnitude estimates as a function of log circle size) ranged from 0.5 to 1.0 with a mean of 0.7 and a standard deviation of 0.1. Similar results (Table M-3) were obtained for separate Trial 2 and Trial 3 slopes on log magnitude estimates as a function of log circle size. The Pearson product-moment correlation coefficient (Linton and Gallo, 1975, pp. 347-352) for Trial 2 slopes and Trial 3 slopes was significant ($r = 0.576$ for $\alpha = 0.05$, df = 10; $r_{\text{observed}} = 0.98$, df = 10). The coefficient of determination (Linton and Gallo, 1975, pp. 344-346) was 0.96, indicating that 96% of the total variance in the subjects' slopes can be accounted for by a linear relationship.

Discussion

The very strong positive correlations between the magnitude estimates for Trial 2 and Trial 3 stimuli suggest good within-session test-retest reliability for all subjects. The very strong positive correlation between the Trial 2 slopes and Trial 3 slopes for the log-log functions also suggest excellent within-session test-retest reliability.

The linear nature of the log-log functions suggests that the data represent a power function (Stevens, 1957). The mean slope of the power

Table M-3. Slopes for the least squares lines of best fit for three functions: (1) log magnitude estimates for Trial 2 as a function of log circle size (in^2), (2) log magnitude estimates for Trial 3 as a function of log circle size (in^2), and (3) log geometric mean magnitude estimates across Trials 2 and 3 as a function of log circle size (in^2).

Subject	Slopes for Trial 2 Data	Slopes for Trial 3 Data	Slopes for Trials 2 and 3 Data
1	0.842	0.871	0.857
2	0.610	0.591	0.600
3	0.986	0.986	0.986
4	0.893	0.936	0.914
5	0.668	0.668	0.668
6	0.513	0.542	0.527
7	0.650	0.680	0.665
8	0.623	0.600	0.612
9	0.601	0.591	0.596
10	0.668	0.639	0.654
11	0.747	0.767	0.757
12	0.728	0.795	0.762
$\overline{X}$	0.711	0.722	0.717
S.D.	0.136	0.147	0.141

function (i.e., 0.7) is the same as the slope reported by Stevens (1957). According to Stevens' (1957) power law, one would expect the subjective magnitude of visual size to increase as the 0.7 power or exponent of the area. In other words, $\psi = \kappa\phi^{0.7}$ where ψ = subjective magnitude, κ = a scaling factor equal to the intercept, and ϕ = stimulus magnitude (i.e., area).

It should be cautioned that the results of this study do not consider reliability over time or the extent, if any, to which the training provided by the newly developed materials was responsible for the results obtained. The results do suggest, however, that after participating in this program, initially naive subjects are capable of performing a visual magnitude estimation task yielding within-session results similar to those reported by Stevens (1975).

APPENDIX N

INSTRUCTIONS AND RESPONSE SHEETS FOR LISTENING TASKS

Please read these instructions as you listen to them. You are going to hear some short speech passages in one ear. We want you to assign numbers to these passages in much the same way you did to the circles and squares. You will write these numbers on a response sheet similar to the one you used before.

Instead of seeing pairs of shapes, you will be hearing pairs of speech passages. The first passage in each pair will always be the standard passage, and the second passage will always be the comparison passage. The quality of the comparison passage will change from time to time.

First, you will hear the word "standard" followed by a single spoken passage. When this happens, you are to assign a number to the passage. The number you pick should represent your impression of the overall quality or "goodness" of the passage. You may pick any number you want. Write the number you pick in the block labeled "S" on your response sheet.

Next, you will hear the pairs of speech passages. Each pair or item will be preceded by a spoken item number which corresponds to a block number on your response sheet. Your job will be to refer to the number you initially gave to the standard and then pick a number for the comparison. The number for the comparison should represent the quality of the comparison passage relative to the quality of the standard. In other words, the number you pick for the comparison should relate to the number for the standard in the same way that the quality of the comparison relates to the quality of the standard.

Let's take an example. Say you hear the word "standard" followed by a speech passage. You assign that passage some number, say 16. Then, you hear an item number followed by the same standard passage and then a comparison passage in that order. If the quality of the comparison sounds twice as good as the standard, the comparison gets the number 32. If the comparison sounds one-fourth as good, you'd call it 4.

After hearing a trial of seven pairs of passages, you will hear a new trial beginning with a new standard passage. Just follow the same procedure used before.

We're almost ready to start. Remember, this is not an intelligence test or a "trick" test. Even though there are no wrong answers, we want you to pay careful attention to what you hear. Remember, your job is to recall the number you gave the standard and to assign a number to the comparison that represents the quality of the comparison relative to the quality of the standard.

Look at your response sheets. After each item you will have about 5 seconds to write a number in the appropriate box. Start with the blocks for Trial 1 in the lefthand column and work from top to bottom as you did before. You should have four response sheets. The first sheet is for practice items and has three columns. The last three sheets are for test items and only have two columns. I will be the talker for the practice items and the test items.

If you want to reread these instructions or ask a question, inform the experimenter. Also, inform the experimenter when you are ready to begin the practice task.

REFERENCES

Page_____ of_____

RESPONSE SHEET

Subject No.____________ Date_____________ Session______________

Stimulus Sets		
Trial 1	Trial 2	Trial 3
S	S	S
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7

Page____ of____

RESPONSE SHEET

Subject No.____________ Date______________ Session____________

Stimulus Sets

Trial 1		Trial 2	
S		S	
1		1	
2		2	
3		3	
4		4	
5		5	
6		6	
7		7	

REFERENCES

AMERICAN NATIONAL STANDARDS INSTITUTE, American National Standard Methods for Electroacoustic Characteristics of Hearing Aids. ANSI S3.3-1960. New York: American National Standards Institute (1960).

AMERICAN NATIONAL STANDARDS INSTITUTE, American National Standard for Calculation of the Articulation Index. ANSI S3.5-1969. New York: American National Standards Institute (1969).

AMERICAN NATIONAL STANDARDS INSTITUTE, American National Standards Specifications for Audiometers. ANSI S3.6-1969. New York: American National Standards Institute (1969).

ANDERSON, H., BARR, B., and WEDENBERG, E. Early diagnosis of the eighth-nerve tumors by acoustic reflex tests. ACTA Otolaryngologica Supplement 263, 232-237 (1970).

BERGER, K. Speech audiometry. In D. Rose (Ed.), Audiological Assessment. Englewood Cliffs, N.J.: Prentice-Hall (1978).

BOCCA, E. and CALEARO, C. Central hearing processes. In J. Jerger (Ed.), Modern Developments in Audiology. New York: Academic Press (1963).

BURNEY, P. A survey of hearing aid evaluation procedures. ASHA, 14, 439-444 (1972).

CARHART, R. and JERGER, J. Preferred method for clinical determination of pure tone thresholds. Journal of Speech and Hearing Disorders, 24, 330-345 (1959).

CHIAL, M.R. The masked intelligibility threshold test for speech communication system evaluation. Doctoral dissertation, University of Wisconsin (1973).

CHIAL, M.R. and DANIEL, S.W. Hearing aid quality judgments by normal and dysacusic listeners. Paper presented at the Annual Convention of the American Speech and Hearing Association, Chicago (1977).

CHIAL, M.R. and HAYES, C.S. Hearing aid evaluation methods: Some underlying assumptions. Journal of Speech and Hearing Disorders, 39, 270-279 (1974).

FANG, I.E. The "easy listening formula". Journal of Broadcasting, XI, 63-68 (1966-1967).

FLETCHER, H. A method of calculating hearing loss for speech from an audiogram. Journal of the Acoustical Society of America, 22, 1-5 (1950).

FLETCHER, H. and GALT, R.H. The perception of speech and its relation to telephones. Journal of the Acoustical Society of America, 22, 89-150 (1950).

FLETCHER, H. and MUNSON, W. A. Loudness, its definition, measurement and calculation. Journal of the Acoustical Society of America, 5, 82-108 (1933).

FRENCH, N. and STEINBERG, J. Factors governing the intelligibility of speech sounds. Journal of the Acoustical Society of America, 19, 90-119 (1947).

GESCHEIDER, G.A. Psychophysics: Method and Theory. Hillsdale, N.J.: Lawrence Erlbaum Associates (1976).

GRAY, T. and SPEAKS, C. Ability of hearing-impaired listeners to understand connected discourse. Journal of the American Audiology Society, 3, 159-166 (1977).

HAYS, W. Statistics for Psychologists. New York: Holt, Rinehart and Winston (1963).

HECKER, M. and GUTTMAN, N. Survey of methods for measuring speech quality. Journal of the Audio Engineering Society, 15, 400-403 (1967).

IEEE. Institute of electronics and electrical engineers recommended practice for speech quality measurements, Number 297. IEEE Transactions on Audio and Electroacoustics, AU-17, 225-246 (1969).

JEFFERS, J. Quality judgment in hearing aid selection. Journal of Speech and Hearing Disorders, 25, 259-266 (1960).

JERGER, J. Clinical experience with impedance audiometry. Archives of Otolaryngology, 92, 311-324 (1970).

JERGER, J. and JERGER, S. Diagnostic significance of PB word functions. Archives of Otolaryngology, 93, 573-580 (1971).

JERGER, J. and JERGER, S. Clinical validity of central auditory tests. Scandinavian Audiology, 4, 147-163 (1975).

JERGER, J. and THELIN, J. Effects of electroacoustic characteristics of hearing aids on speech understanding. Bulletin of Prosthetics Research, 10, 159-197 (1968).

KASTEN, R. and REVOILE, S. Variability of electroacoustic characteristics of hearing aids. Paper presented at the Annual Convention of the American Speech and Hearing Association, Chicago (1965).

KATZ, J. The use of staggered spondaic words for assessing the integrity of the central auditory nervous system. Journal of Auditory Research, 2, 327-337 (1962).

KERST, S.M. MAGEST: A FORTRAN program for analyzing magnitude estimation data. Behavior Research Methods and Instrumentation, 10, 737 (1978).

KIRK, R.E. Experimental Design: Procedures for the Behavioral Sciences. Monterey, CA: Brooks/Cole (1968).

KLARE, G.R. The Measurement of Readability. Ames, Iowa: The Iowa State University Press (1963).

LASSMAN, F. A noise-interaction theory of intelligibility and its relation to hearing-aid amplification. Auditory Rehabilitation in Adults, Seminar Proceeding, Cleveland, June 8-12 (1964).

LICKLIDER, J.C.R. Effects of amplitude distortion upon the intelligibility of speech. Journal of the Acoustical Society of America, 18, 429-434 (1946).

LICKLIDER, J.C.R. and HELD, R. Effects of various types of nonlinear distortion upon the intelligibility of speech. Journal of the Acoustical Society of America, 24, 114 (1952).

LINTON, M. and GALLO, P.S. The Practical Statistician: Simplified Handbook of Statistics. Monterey, CA: Brooks/Cole (1975).

MARKS, L.E. Sensory Processes: The New Psychophysics. New York: Academic Press (1974).

MARTIN, F.N. and FORBIS, N.K. The present status of audiometric practice: A follow-up study. ASHA, 12, 531-541 (1978).

MARTIN, F.N. and PENNINGTON, C.D. Current trends in audiometric practices. ASHA, 13, 671-677 (1971).

MCGEE, V.E. Determining perceptual spaces for the quality of filtered speech. Journal of Speech and Hearing Research, 8, 23-38 (1965).

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

MUNSON, W. and KARLIN, J. Isopreference method for evaluating speech-transmission circuits. Journal of the Acoustical Society of America, 27, 1213-1219 (1962).

OCHIAI, Y. and FUKUMURA, T. Timbre study of vocalic voices. Memoirs of the Faculty of Engineering at Nagoya University, 5, 253-280 (1953).

OCHIAI, Y. and FUKUMURA, T. Timbre study of vocalic voices viewed from subjective phonal aspect. Parts I, II, and III. Memoirs of the Faculty of Engineering at Nagoya University, 8, 1-18, 203-221, 222-239 (1956).

OLSON, W.O. and NOFFSINGER, D. Comparison of one new and three old tests of auditory adaptation. Archives of Otolaryngology, 99, 94 (1974).

OYER, H.J. and FRANKMANN, J.P. The Aural Rehabilitation Process, A Conceptual Framework Analysis. New York: Holt, Rinehart and Winston (1975).

POWERS, G.L. and SPEAKS, C. Intelligibility of temporally interrupted speech. Journal of the Acoustical Society of America, 54, 661-667 (1973).

PUNCH, J.L. Quality judgments of hearing aid processed speech and music by normal and otophathologic listeners. Journal of the American Auditory Society, 3, 179-188 (1978).

PUNCH, J.L. and BECK, E.L. Aided speech quality judgments: Effects of varying low-cutoff frequency. Paper presented at the Annual Convention of the American Speech and Hearing Association, Atlanta (1979).

PUNCH, J.L. and CIECHANOWSKI, J.M. Reliability of paired-comparison quality judgments in hearing aid evaluation. Paper presented at the Annual Convention of the American Speech and Hearing Association, Chicago (1977).

PUNCH, J.L. and PARKER, D.A. Validity of paired-comparison listener preferences in hearing aid evaluation. Paper presented at the Annual Convention of the American Speech and Hearing Association, Atlanta (1979).

ROTHAUSER, E.H., URBANEK, G.E. and PACHL, W.P. Isopreference method for speech evaluation. Journal of the Acoustical Society of America, 44, 408-418 (1968).

ROTHAUSER, E.H., URBANEK, G.E. and PACHL, W.P. A comparison of preference measurement methods. Journal of the Acoustical Society of America, 49, 1297-1308 (1971).

SCHWARTZ, D.M., WALDEN, B.E. and PROSEK, R.A. Electroacoustic correlates of hearing aid quality judgments. Paper presented at the Annual Convention of the American Speech and Hearing Association, Atlanta (1979).

SHANNON, C.E. and WEAVER, W. The Mathematical Theory of Communication. Urbana: University of Illinois Press (1949, 1963).

SMALDINO, J. The differentiation of low fidelity circuitry by behavioral response. Doctoral dissertation, University of Florida (1974).

SPEAKS, C. and TROOIEN, T.T. Interaural alternation and speech intelligibility. Journal of the Acoustical Society of America, 56, 640-644 (1974).

STEVENS, J.C. and TULVING, E. Estimations of loudness by a group of untrained observers. American Journal of Psychology, 70, 600-605 (1957).

STEVENS, S.S. A scale for psychological magnitude: Loudness. Psychological Review, 43, 405-416 (1936).

STEVENS, S.S. On the psychophysical law. Psychological Review, 64, 153-181 (1957).

STEVENS, S.S. Power-group transformations under glare, masking, and recruitment. Journal of the Acoustical Society of America, 39, 625-735 (1966).

STEVENS, S.S. *Psychophysics: Introduction to Its Perceptual, Neural and Social Prospects*. New York: John Wiley and Sons (1975).

STEVENS, S.S. and GALANTER, E.H. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, *54*, 377-411 (1957).

STEVENS, S.S. and POULTON, E.C. The estimation of loudness by unpracticed observers. *Journal of Experimental Psychology*, *51*, 71-78 (1956).

THALMANN, R. Cross-modality matching in a study of abnormal loudness functions. *Laryngoscope*, *75*, 1708-1726 (1965).

WELDELE, F.J. The usefulness of psychophysical listener judgments in hearing aid evaluation. Masters thesis, Kent State University (1973).

WELDELE, F.J. and MILLIN, J.P. The usefulness of qualitative listener judgments in evaluating hearing aids. *Ohio Journal of Speech and Hearing*, *11*, 39-47 (1975).

WILDER, H.B., LUDHUM, R.P., and BROWN, H.M. *This is America's Story*. Boston: Houghton, Mifflin Company (1954).

WILSON, R.H., COLEY, K.E., HAENEL, J.L., and BROWNING, K.M. Northwestern University Auditory Test No. 6: normative and comparative intelligibility functions. *Journal of the American Auditory Society*, *1*, 221-228 (1976).

WINER, B.J. *Statistical Principles in Experimental Design*. New York: McGraw-Hill Book Company (1971).

WITTER, H. and GOLDSTEIN, D. Quality judgments of hearing aid transduced speech. *Journal of Speech and Hearing Research*, *14*, 312-322 (1971).

YONOVITZ, A., BICKFORD, B.J., LOZAR, J. and FERRELL, D.R. Electroacoustic distortions: Multidimensional analysis of hearing aid transduced speech and music. Paper presented at the 1978 IEEE International Conference on Acoustics, Speech and Signal Processing, Tulsa (1978).

ZERLIN, S. A new approach to hearing-aid selection. *Journal of Speech and Hearing Research*, *5*, 370-376 (1962).