INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND

8103947

MARGULIES, MARIS KABAT

EFFECTS OF TALKER DIFFERENCES ON SPEECH INTELLIGIBILITY IN THE HEARING IMPAIRED

City University of New York PH.D. 1980

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1980

by

Margulies, Maris Kabat

All Rights Reserved

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs _____
2. Colored illustrations _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Print shows through as there is text on both sides of page _____
6. Indistinct, broken or small print on several pages ✓
7. Tightly bound copy with print lost in spine _____
8. Computer printout pages with indistinct print _____
9. Page(s) _____ lacking when material received, and not available from school or author
10. Page(s) _____ seem to be missing in numbering only as text follows
11. Poor carbon copy _____
12. Not original copy, several pages with blurred type _____
13. Appendix pages are poor copy _____
14. Original copy with light type _____
15. Curling and wrinkled pages _____
16. Other _____

University Microfilms International

300 N. ZEEB RD., ANN ARBOR, MI 48106 (313) 761-4700

EFFECTS OF TALKER DIFFERENCES ON SPEECH INTELLIGIBILITY IN THE HEARING IMPAIRED

by

MARIS KABAT MARGULIES

A dissertation submitted to the Graduate Faculty in Speech and Hearing Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1980

© COPYRIGHT BY

MARIS KABAT MARGULIES

1980

This manuscript has been read and accepted for the Graduate Faculty in Hearing and Speech Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

August 16, 1980
date

Chairman of Examining Committee

August 14, 1980
date

Executive Officer

Dr. Thomas Fay

Dr. Irving Hochberg

Dr. Gerald Studebaker
Supervisory Committee

The City University of New York

Abstract

EFFECTS OF TALKER DIFFERENCES ON SPEECH INTELLIGIBILITY IN THE HEARING IMPAIRED

by

Maris Kabat Margulies

Adviser: Harry Levitt

Five groups of ten listeners each (normal hearing, conductively impaired, flat-, moderately sloping-, and severely sloping sensori-neurally impaired) were presented with recordings of the CID sentences of everyday speech (CHABA) spoken by ten different talkers (five male and five female). In addition, recordings of nonsense syllables for two talkers (one male and one female from the above group) were also used. Five types of background noise were added to the sentence materials: high-pass, low-pass, white, male-spectrum and female-spectrum shaped noise. The ratio of speech peaks to rms noise level was set at 0 dB for the sentence materials. Four experimental conditions were used for the nonsense syllable materials: frequency filtering with a slope of 0 dB/octave, +6 dB/octave and -6 dB/octave (all in quiet) and 0 dB/octave slope condition in noise at a S/N ratio of 0 dB. In addition to obtaining

intelligibility scores, subjects were required to rate the intelligibility of each talker on a scale of 1-7.

Significant differences in intelligibility were obtained between talker gender, experimental groups, experimental conditions (added background noise and/or frequency filtering), and listener age. On the sentence materials the female talkers, on the average, were more intelligible than the male (for a common speech-to-noise ratio). There was, however, a substantial overlap between the intelligibility scores of individual male and female talkers. For the nonsense syllable materials, where only one talker of each gender was used, the male talker yielded higher discrimination scores than the female. This male talker happened to be the most intelligible of the male talkers used in the study, whereas the female talker was of average intelligibility. No significant differences were observed between male and female listeners.

As expected, intelligibility scores decreased with increase in degree of hearing impairment but the relative masking and filtering effects varied between listener groups and talker gender. Younger listeners scored higher than older listeners except in the conductive hearing loss group. This apparent exception is believed to be the result of a very high score shown by one listener.

The intelligibility rankings showed a higher degree of consistency with the objective measure of intelligibility for the normal

hearing listeners. The consistency between the rankings and measured intelligibility decreased as the degree of hearing impairment, especially in the higher frequencies, increased. There was no significant male-female talker difference in the consistency of the ratings.

DEDICATION

To my Parents:

My Father for his love and support, caring and helping always;

My Mother, for had she the opportunity, this success would be hers.

I love you both.

ACKNOWLEDGEMENT

This last portion of a dissertation to be written is perhaps the most difficult. It is evident that the study could never have been completed without the efforts of many people who deserve recognition for their time, energy, insight and concern.

In this case my committee must be commended for the many years of effort involved in bringing this project to completion. Professors Harry Levitt, Irving Hochberg and Gerald Studebaker were always available, knowledgeable and willing to listen to my successes and failures.

Cecil Redmond in the Communication Sciences Laboratory and Ron Slosberg and Harvey Stromberg on the computer were invaluable as were Lyn Reid and Gertie Glatt in the department office.

The staff and patients of the Barnert Memorial Hospital deserve thanks for having seen the merit of my efforts and allowing me the use of the facilities and their time. Jackie Edelstein, my friend as well as my supervisor, commiserated with those subjects who sat and listened to sentences and seemingly endless lists of nonsense syllables.

My dear husband, Lee, my daughters, and the rest of my

family accepted my absences and understood my extended trips to the attic. It's good to be able to come downstairs.

To those friends who did not see me, you were not forgotten; nor were those babysitters, especially Harriet Schwartz, who cared for my girls as though they were their own.

There is one special friend, though, who deserves a special accolade. He was a supporter who never gave up, even when I thought I would. A very special thank you to my advisor, chairman of my committee and special person, the tenacious "Dr. Smith" who always managed to know when he was needed. Thank you, Harry Levitt.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

It has frequently been reported that for hearing-impaired persons, female speech is not as intelligible as male speech (Grubb, 1963). This notion that hard of hearing individuals hear the voices of men "better" or "easier" than the voices of women is a common informal observation in many clinical situations (Palmer, 1955). Data supporting this observation may be found in the MedResCo Report (Medical Research Council, 1947) in which speech reception scores for female talkers were consistently poorer than those for male talkers. A similar result was obtained by Black and Levitt (1969) who, in an investigation of the effects of frequency response on speech intelligibility for hearing-impaired listeners, found higher intelligibility scores for the male talker than for the female.

It is important to determine if indeed female talkers are less intelligible than male talkers, and if so, possible reasons why. This is especially important since there are more female audiologists than male audiologists practicing (personal communica-

tion, ASHA, 1975)[1] and surveys have shown that live-voice test procedures are used 64.4% of the time (Martin and Pennington, 1971).

Since the early development of speech audiometric materials at the Harvard Psycho-Acoustic Laboratories (Egan, 1948; Hudgins, Hawkins, Karlin & Stevens, 1947) and their later refinement at the Central Institute for the Deaf (Hirsh, Davis, Silverman, Reynolds, Eldert & Benson, 1952), speech audiometry has become a well established diagnostic tool for assessing the speech function. In the usual evaluation these tests are primarily of two types: speech reception threshold (SRT) tests and tests of speech discrimination. The first of these is a measure of the person's sensitivity for speech. The other is a suprathreshold task which is used to provide an estimate of the individual's degree of difficulty in understanding speech. The stimuli may vary and may be delivered from tape or phonograph recordings or be monitored-live-voice (MLV).

While the most commonly used tests for measuring the speech function are the CID W-1 and W-22 lists (Martin and Pennington, 1971), and since these are performed predominantly by live-voice procedure, it is necessary to consider the effects of between-talker differences on the test materials. It has been shown

---

[1]These data were obtained at the beginning of the study.

that the effects of talker change are substantial, although they vary according to the material used (Hirsh et al., 1954; Palmer, 1955; Harris et al., 1960; Weiss, 1968; Kruel, 1968, 1969; Pascoe, 1974, 1975).

The present study is an attempt to investigate the differences in the intelligibility of male and female voices as heard by hearing impaired listeners. It is concerned specifically with an examination of speech discrimination ability of hearing impaired listeners for male and female talkers for various test conditions and test materials.

CHAPTER II

REVIEW OF THE LITERATURE

Methods of Measuring Speech Perception

Historical perspective

Tests of a hearing-impaired individual's ability to understand speech have existed for as long as there has been an interest in the effects of hearing impairment on speech understanding. Originally, assessments of understanding were accomplished by simply speaking to the person suspected of having a hearing impairment and making a judgment about his ability to understand. The listener's performance was compared to that of persons unaffected by auditory impairments (Hirsh, 1947). Such a crude method did not allow for calibration or control of the signal with respect to intensity or message content.

Bryant (1904) attempted to overcome this difficulty through a system he designed which controlled the level of the speech signal by means of a phonograph and valve attachment. Campbell (1910) developed the first analytic approach to the measurement of speech intelligibility by asking the listener to repeat a list of nonsense

syllables which had been read over a telephone channel. The percentage of correct identification was the measure of relative intelligibility.

It was more than 20 years after Bryant (1904) advocated a recorded presentation of material that the use of a recorded auditory test became available. This was not a test for assessing speech understanding ability, but for determining an individual's hearing loss for speech, and used numbers as test stimuli (Fletcher, 1929). Continued efforts were made to refine these procedures, notably efforts by French & Steinberg (1929) who also developed word and sentence lists of interrogative and imperative sentences for intelligibility testing as part of their research at Bell Telephone Laboratories.

Development of materials

The earliest tests of speech understanding were used to rate speech transmission systems and were referred to as articulation tests. The resulting score was an articulation score. By plotting a graph with percentage of items correct on the ordinate and intensity on the abscissa, one can generate an articulation-gain function. Such a graph depicts the relationship of understanding to changes in intensity for the particular stimulus material. The articulation-gain function has been shown to vary depending upon the type of material used (Fletcher & Steinberg, 1929; Hirsh, et al.,

1952; Hudgins, et al., 1947; Miller, Heise, & Lichten, 1951; Traul & Black, 1965). In general, the more information there is in an utterance, the steeper is its articulation-gain function. Three-syllable words typically contain more information than two-syllable words which typically contain more information than monosyllabic words, which in turn contain more information than nonsense syllables. Correspondingly, the articulation-gain functions for two-syllable words are steeper than those for monosyllables which in turn are steeper than nonsense syllables (Hirsh, 1952). Similarly, the articulation-gain functions for sentences are very steep.

In addition to the syllable articulation lists, Fletcher & Steinberg (1929) also devised the sentence lists mentioned earlier. Since the purpose of these tests was to test the accuracy of perception, each sentence was designed to convey a simple idea. A response was considered correct if the listener could either record the sentence or answer it correctly.

There were two major criticisms of the Bell Telephone Laboratories sentence intelligibility lists (Egan, 1948; Hirsh, 1952). It was pointed out that the questions often required specific knowledge of New York, and additionally, the questions often were beyond the ability of many people to answer correctly.

A major contribution to the development of speech testing procedures was the extensive battery of tests developed by the

Harvard Psychoacoustics Laboratory (PAL) during World War II. These tests were developed initially for the evaluation of military speech communication systems, but the broader more useful applications of the tests were soon recognized. While the battery initially included nonsense syllable lists, phonetically balanced monosyllabic word lists, spondaic word lists, and phonetically balanced sentence lists, it was the monsyllabic and spondaic word lists which found widespread clinical application.

Phonetically balanced word lists (in abbreviated form referred to as PB word lists) were devised to meet the following criteria: (1) monosyllabic structure, (2) equal average difficulty, (3) equal range of difficulty, (4) equal phonetic composition, (5) a composition representative of English speech, and (6) words in common usage (Egan, 1948).

The early clinical application of the PAL materials met with a number of practical problems. Many patients had difficulty with the vocabulary which contained a number of unfamiliar words. Also, the speaker on the recordings was Rush Hughes, a professional radio announcer, who recorded the lists without a carrier phrase. Eldert & Davis (1951) reported on the clinical use of these recordings and indicated that chief among the problems were low reliability, a wide range of difficulty among the individual test items, and an uneven distribution of difficulty among the eight lists.

Hirsh et al. (1952) reported that clinical use of PAL Auditory Tests No. 9 and 12 and of the PB word lists (both recorded and live-voice versions) revealed deficiencies which needed to be corrected: different thresholds were obtained from different recordings of the spondee lists and the vocabulary of the PB word lists was too unfamiliar for many patients.

In an attempt to remedy this situation, the Harvard monosyllabic and spondaic word lists were revised systematically for speech audiometry. The phonetic balance of the mono-syllabic word lists was maintained, but only familiar words were used and standardized recordings of several different randomizations were prepared (Hirsh, et al., 1952). According to Martin & Pennington (1971), these tests, known as the CID Auditory Tests W-1, W-2 (to measure Speech Reception Threshold), and W-22 (to measure speech discrimination using monosyllabic words) are the most commonly used tests in speech audiometry.

Materials available for discrimination testing

Several methods have been advanced for the measurement of speech intelligibility. Generally, these include testing with single syllables, meaningful isolated words, or meaningful phrases and sentences with their contextual information. The most common word materials have already been discussed. The Harvard sentence

lists were also considered inappropriate for clinical use because their syntactic structure was atypical and they contained many uncommon words. The sentences were all approximately the same length and there was little variation in form between the sentences. Following criteria formulated by CHABA, researchers at CID developed sentence lists that would be typical of everyday conversational speech in terms of vocabulary, sentence structure, and sentence length (Silverman & Hirsh, 1955). These sentences are generally known as the CHABA sentences, or CID sentences of everyday speech.

Tests of speech intelligibility using a closed-response-set format are among the newest materials available for clinical use. In this type of test the subject is restricted to a specified set of response alternatives. Since the subject can be trained with each set of response foils before testing, learning and practice effects are reduced considerably and the tests can be used repeatedly. There are two important advantages to the use of this type of material: (1) the error patterns can be studied to provide diagnostic information, and (2) the researcher can tailor the choice of response foils to meet specific test needs.

The closed-set paradigm for word discrimination followed the development of a rhyme test developed by Fairbanks (1958). House, et al. (1965), developed the Modified Rhyme Test in which the subject is supplied with a response set of six words, each differ-

ent from the target word by a single phonetic feature. Fifty sets of items are presented. Variations on the rhyme test procedure have been proposed by Kreul, et al. (1968). As yet, however, conventional clinical evaluations do not make use of the closed-response-set format except in the case of picture tests used with children.

An attempt to make use of a closed-response-set format using sentence materials is the Synthetic Sentence Identification (SSI) Test developed by Speaks and Jerger (1965). Kalikow, Stevens, & Elliot (1977) have developed a test which measures the effect of context on speech intelligibility. The Speech Perception in Noise Test (SPIN) tests the discrimination of each test word in two contexts, one where the likelihood of occurrence is high, and one in which the word is not likely to occur. The differences between the two scores provide a quantitative measure of the effect of word predictability on discrimination.

Three phases in the development of speech intelligibility materials have been identified: (1) the very early work of the 1920-1930's, in which the need was recognized, but only the beginnings of practical solutions were available; (2) the development and refinement of the procedures produced by the Harvard Psychoacoustics Laboratory, and (3) some of the newest additions to testing, including the closed-response-set paradigm.

Effects of Masking on Speech Intelligibility

Many experiments have measured the relationship of speech discrimination to various types of background interference. Speech (Miller & Licklider, 1950; Carhart, Tillman, & Johnson, 1966), continuous noise (Miller & Garner, 1948), interrupted noise (Miller & Licklider, 1950; Miller & Garner, 1948), modulated noise (Carhart, Tillman, & Johnson, 1966; Wilson & Carhart, 1969), and other complex sounds have been used as noise background (Johnson & Young, 1974; Kacena & Nicholls, 1974). In summary, these studies demonstrated that intelligibility decreased as the level of background noise increased. Miller & Garner (1948), using interrupted bursts of white noise as interference, found an increase in duration of the interruption improved the intelligibility. Pollack (1955) studied the intelligibility of connected speech against a background of modulated white noise and found that varying the signal-to-noise ratio at rather slow rates provided intelligibility comparable to a steady-state signal-to-noise ratio at the average level. In addition, it was found that varying the absolute levels but keeping the signal-to-noise ratio constant had no appreciable effect on measured speech intelligibility.

These experiments dealt with normal-hearing subjects. Ross et al. (1965) using a continuous noise as the interfering signal, reported the effects of noise were equal for subjects with normal and abnormal hearing. This result is not supported by other research.

Palva (1955) measured speech thresholds and discrimination in the presence of continuous noise. He found that discrimination was not affected significantly in normal-hearing listeners and those with a conductive hearing loss but was significantly reduced in individuals with sensori-neural hearing loss. He found no difference in the effect of noise on the speech reception thresholds of the three groups. Cooper and Cutts (1971), using ambient noise (recorded in a cafeteria) found that discrimination scores were significantly poorer for individuals with sensori-neural hearing losses than for normal-hearing persons. Measuring discrimination for mono-syllables against a background of competing sentences, Carhart & Tillman (1970) found that subjects with conductive hearing loss performed the same as normal-hearing subjects. Persons with sensori-neural loss were more severely handicapped by competing messages. Wilson & Carhart (1969), using spondee words, compared the influence of pulsed and modulated noise on the hearing thresholds of normal-hearing subjects to thresholds of subjects with diagnosed cochlear otosclerosis. They found that subjects with the sensori-neural hearing loss performed significantly worse than subjects with normal hearing.

Dirks et al. (1969), using normal-hearing subjects, showed that the type of speech material presented is an important variable when testing discrimination in a noisy environment. Apparently the

amount of masking is the complex result of an interaction between the type of speech-test material, the rate of modulation or interruption, and the status of the auditory system of the listener, as well as the more commonly considered variables of masker spectrum and speech-to-noise ratio (Miller, 1947).

Shapiro, Melnick & VerMeulen (1972) studied the effects of modulated noise on speech intelligibility of people with sensori-neural hearing loss, and found their data consistent with past research in that under all the experimental noise conditions, normal-hearing subjects performed better than those with sensori-neural hearing loss. These findings are consistent with the more recent research using a closed-response-set format (Levitt et al., 1978), and proposals for application of a sentence test utilizing the difference in discrimination ability of hearing-impaired listeners in quiet and in noise (Kalikow, Stevens and Elliot, 1977). In a recent paper presented by Gengel & Marks (1980), it was reported that not only hearing loss but age must be considered when investigating the discrimination ability of hearing-impaired listeners in noise. They found that not only was the average performance of the older listeners below that of the younger listeners, but that there was an age-talker interaction as well.

Talker Effects on Speech Intelligibility

There are many tests which can be used to assess speech discrimination ability in the hearing-impaired. Unfortunately, the relationship between the degree of hearing loss and word discrimination ability is not clear, especially in cases of sensori-neural hearing loss.

Carhart (1965) stated that the test results which were obtained by different talkers are not comparable "unless the talkers have been demonstrated to be equivalent." He also noted that, where this information is not available, comparisons can be made only on the results obtained by a single talker. Brandy (1966), however, has demonstrated that a given talker will vary his presentation of the same words on successive days. Using recorded presentations, which were controlled for variations in intensity, substantially less variability was obtained in listeners scores as compared to live-voice presentations.

Harris, Haines, Kelsey, and Clack (1960) reported inter-talker effects in their study on the effects of low-fidelity circuitry on speech intelligibility. They reported that scores varied significantly between talkers for identical experimental conditions. House, Williams, Hecker and Kryter (1965) in their study dealing with the development of the Modified Rhyme Test, found a significant difference in intelligibility between their talkers. Kreul, Nixon, Kryter,

Bell, Lang and Schubert (1968) and Kreul, Bell and Nixon (1969), studied the effects of several variables on the reliability of discrimination scores using normal-hearing subjects. One of the variables studied was the effect of different talkers. They reported that the test-retest scores for either of their talkers on different sessions were not significantly different. The overall level of difficulty reportedly remained stable over several utterances. It should be noted that these findings were obtained using recorded materials and although there were differences in intelligibility scores for the different talkers, these talkers were consistent, i.e. their talkers were not equally intelligible, but by using recorded materials they obtained low test-retest variability for each talker.

Beyer, Webster and Dague (1969) agree with this observation. They found differences in the clinical version of the MRHT to be related to talker differences, with one of the talkers consistently producing lower intelligibility scores than the other two talkers.

The problem of high test-retest variability using live-voice rather than recorded materials has been stressed by Goetzinger (1972) and Carhart (1965). Further, Tillman and Olsen (1973) point out that a standardized test is not possible unless recordings are used. They attribute this to talker differences and state that "...because of talker differences tests administered via monitored-live-voice defy standardization."

Witter and Goldstein (1971) recorded the speech of a male and female talker through hearing aids and had paired comparison judgments made by normal-hearing listeners. They found that hearing aid choices varied for the male and female talkers although the rankings were similar. Bell and Kreul (1972) found, on re-evaluating the reliability of the MRHT that there were again differences between talkers. They had two male and one female talker for the study. The male talkers were both the best and worst talkers in terms of intelligibility, with the female talker midway between the other two talkers.

Pascoe (1974; 1975) recorded the high-frequency word lists by a male and female talker. He found that the test scores were consistently higher for the male talker as compared to the female talker, and that the subjects who had better hearing were less affected than the others by this difference.

Levitt & Resnick (1978) developed a nonsense syllable test for measuring speech intelligibility. They recorded both a male and female version of the test and found the male talker to provide higher intelligibility scores. Mayer (1980) compared sentence discrimination ability for elderly subjects for groups of male and female talkers and found the female talkers to be more intelligible, on average, than the male talkers. It is important to note that all of Mayer's tests have speech against a background of noise, while all

the other studies have been performed in quiet. This difference may account for the higher female intelligibility obtained in Mayer's data.

Acoustic Characteristics of Male and Female Voices

Since there seem to be differences in the perception of male and female speech, the acoustic characteristics that make up this difference must be investigated. The classic data on speech levels and on the shape of the spectrum for male and female talkers are those of Dunn & White (1940). These researchers provided the first relatively comprehensive set of data on both narrow-band and broad-band measurements of speech levels for both male and female talkers. They used six male and five female talkers who were asked to read continuously at a normal voice level, in quiet, from stories containing large proportions of conversation. Dunn & White reported differences between individuals amounting to as much as 18 dB in some bands, and indicated that there were also differences in the shape of the spectrum. They also felt that their sample was too small in size to provide a truly representative result because they could not be sure they had discovered the most extreme individuals.

Subsequent studies, such as those of French & Steinberg (1947) and Benson and Hirsh (1953) replicated these early measurements over wider conditions. Their findings are similar but not identical to Dunn & White.

From these studies come much of the data still considered as representative of speech level and spectrum shape. It should be noted, however, that these values were obtained for talkers in quiet, considered to be equally motivated. Increasing the level of background noise typically results in the talker raising his/her voice. The increase in vocal effort is usually not sufficient to completely override the increase in noise level. Korn (1954) undertook to quantify the relationship between the ambient noise levels and the measured speech levels. Korn found, with conversational vocal efforts, about a 3.5 dB increase in voice level for each 10 dB increase in room noise, whereas Webster & Klumpp (1963) found as much as a 7 dB increase for the same increase in background noise. Kryter (1950) and Pickett (1958) found about a 3 dB increase in shouted vocal effort with a 10 dB increase in masking noise. The results of these studies and a more recent study by Gardner (1966) are consistent, and indicate that the talkers adjust their vocal effort on the basis of how much effort they believe is required for their speech to be understood (Kryter, 1970).

Pearsons, Bennett, & Fidell (1977) undertook a study to provide more information on speech levels in terms of the characteristics of actual listening environments, rather than in controlled environments. They provide a comparison of the speech levels in the categories of casual, normal, raised, loud and shout for male and

female talkers. A comparison for the male and female talkers under the first three conditions showed about a 2 dB difference between the male and female talkers. This difference was found to be larger for the loud and shout categories where the measured differences in speech level between the male and female talkers were 5 and 7 dB, respectively. These data also showed that the shape of the speech spectrum changed in a systematic way, providing more relative power at the higher-frequencies for increased voice level. What is important to consider, however, is that for the normal speech level, the spectra are consistent with previous research.

Kryter (1970) provides an idealized spectrum for male and female voices, as measured at a distance of 1 meter from the lips in a quiet room. The overall level of this male voice is shown to be about 3 dB higher than that of the female voice, but the idealized spectrum of the female voice shows more relative speech power in the higher frequencies (i.e. a more gradual slope in the high frequency region of the speech spectrum of the female talker).

Mayer (1980) obtained the speech spectra of four male and four female talkers as part of her study. She indicates that, as expected, the female talkers show more relative power in the high frequency region.

Concluding Comment

It is apparent from the preceding review of the literature that, from the beginning, researchers and clinicians have been aware of and commented on the variability which exists within and among talkers. Investigations have shown that talker-stimuli and talker-distortion interaction cause significant variations to occur in intelligibility scores. Although the words used may be identical, the acoustic output can be markedly different. As noted by French & Steinberg, "...there are large natural differences between the intensities of the same sound spoken by different people or by the same people at different times."

The review also shows that there are differences in speech intelligibility reported for male and female talkers. For the most part, the female appears to be less intelligibile for the hearing-impaired individual when heard in quiet. The case of male and female for the normal-hearing in noise was mixed. There is a need to investigate this further in view of these inconsistencies and because most intelligibility tests are performed by female clinicians using live-voice procedures. It is important to determine whether this has an effect on the hearing-impaired listener and whether the gender of the talker affects intelligibility differently for the hearing-impaired listener.

CHAPTER III

DESIGN OF THE STUDY

This study was designed to examine the effects of male-female talker differences on the syllable and sentence discrimination scores of listeners with normal hearing, conductive-type hearing losses and a variety of sensori-neural impairments. In order to eliminate the inherent variability which has been shown to exist within a single talker (Brandy, 1966), tape recordings rather than live-voice mode of presentation were used throughout the project.

Subjects

The five subject groups, each composed of ten listeners, were tested at the Hearing and Speech Center of the Barnert Memorial Hospital Center in Paterson, New Jersey. The five subject groups were composed of normal-hearing listeners, listeners with conductive hearing impairments, flat sensori-neural hearing losses, moderately sloping sensori-neural hearing losses and severely sloping sensori-neural hearing losses. Each of the four hearing-impaired groups were subdivided into male and female listeners and also into older and younger listeners. Every attempt

was made to select subjects so as to form balanced sets within each group, but this was not always possible with the available subject pool. A description of each subject group follows with a more detailed individual description in Appendix A. Figure 1 presents a composite audiogram of the mean threshold levels for each experimental group.

Normal-hearing listeners

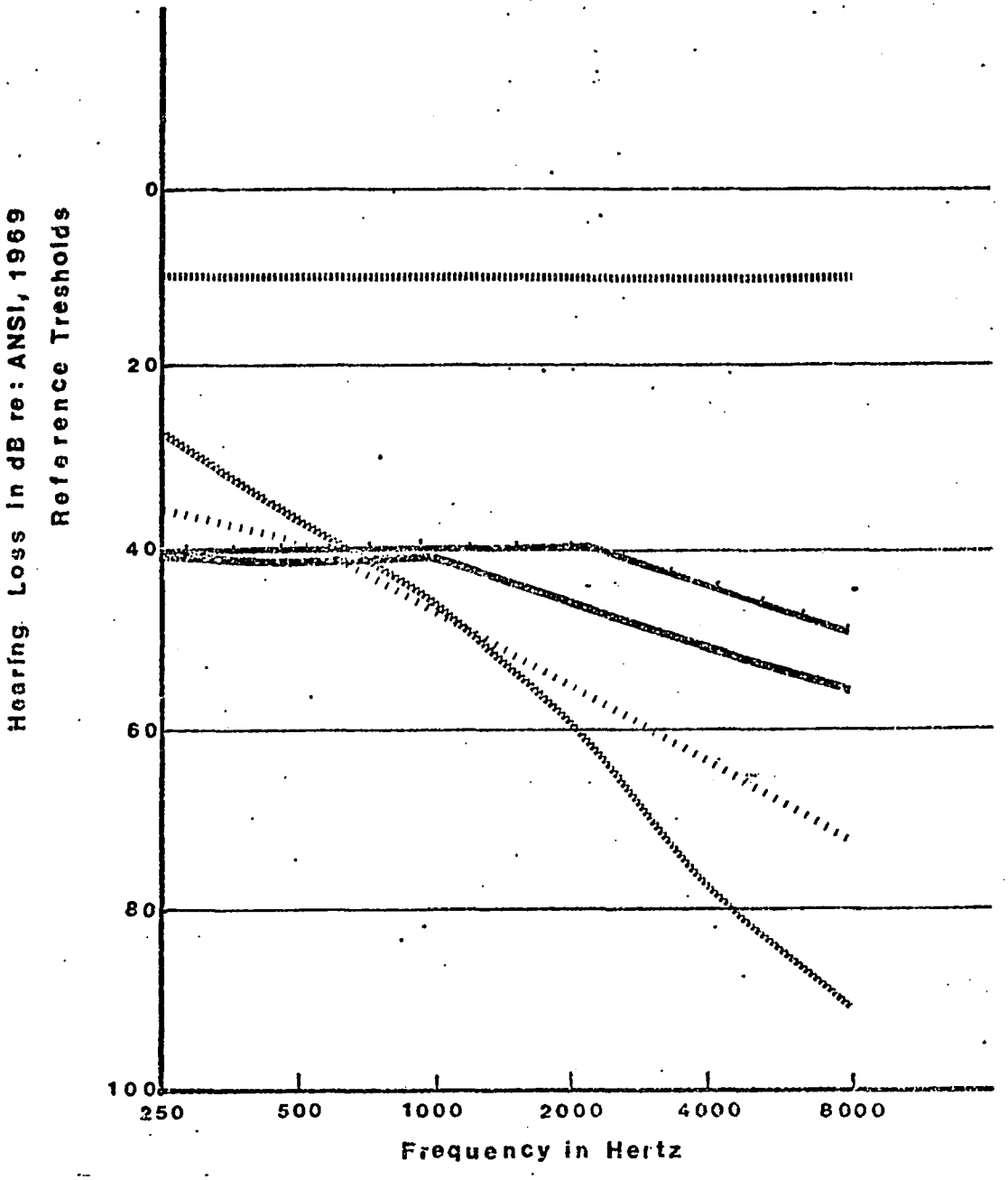
The normal hearing group for the study was composed of ten 20-29 year old adults (8 female and 2 male) who did not have a reported history of ear or hearing problems and who passed a preliminary hearing screening test (pure-tone thresholds were within 10 dB HTL over the 250-8000 Hz range). The subjects also had no prior knowledge of any of the test materials.

Hearing impaired listeners

All potential subjects with hearing impairments were obtained from the recent files of the Barnert Memorial Hospital Hearing and Speech Center. They had all been tested no earlier than 1976 and were re-evaluated for the study. The four groups of listeners with hearing impairment were:

Conductive hearing loss. There were ten subjects who presented a conductive hearing loss, five male and five female listeners

Fig. 1. Composite audiogram of mean pure tone thresholds for each experimental group.


Hearing Loss in dB re: ANSI, 1969
Reference Tresholds
0
20
40
60
80
100
250
500
1000
2000
4000
8000
Frequency in Hertz
Normal-Hearing Listeners
Conductive Hearing Loss
Flat Sensory-Neural Hearing Loss
Moderately Sloping Sensory-Neural Hearing Loss
Severely Sloping Sensory-Neural Hearing Loss

with a median age of 53 years. The median age was used to divide the older from the younger subjects in the experimental design. Only three members of the group had any hearing aid experience and in no instance was a surgically treated ear used for data collection.

Sensori-neural hearing loss. The subjects were thirty sensori-neural hearing impaired adults. They ranged in age from 21-80 years. The mean age was 53.6. There were 15 males and 15 females. The criteria for subject selection were: (1) a relatively flat pure tone audiogram with essentially similar thresholds for the frequencies 250-8000 Hz, (2) moderately sloping audiometric configurations with a 5-10 dB per octave slope, and (3) a more severely sloping audiometric configuration of 15-20 dB per octave.

Preparation of Stimulus Material

Two types of stimulus material were used in the study. The first type was the CID sentences of everyday speech (Davis & Silverman, 1978), developed by a working group of the Armed-Forces National Research Council Committee on Hearing and Bio-Acoustics (CHABA). These CHABA sentences were tape recorded at the City University of New York (CUNY) Communications Sciences Laboratory by ten talkers (5 male and 5 female) and were made available for this study.

Nonsense syllables were the second type of material used.

Recordings of a Nonsense Syllable Test (NST) developed by Levitt and Resnick (1978) were available for both a male and a female talker. These two talkers were among the ten who had recorded the sentence lists.

Sentence materials

The ten CHABA sentence lists were each assigned to one of the ten talkers. The specific list assignments used in the study are shown in Appendix B. A set of test tapes was assembled from the recordings available. The recordings of highest quality in terms of background noise and clarity of articulation were chosen by the experimenter subject to the constraint that each talker be heard on only one list.

The speech level for each recorded list was obtained using a graphic level recorder. The procedure used is described in the next section. The recordings were dubbed onto a test tape such that the average speech level for each sentence list was equal. A 1000 Hz calibration tone at this level was inserted at the start of the test tape.

Noise materials

Five types of noise were used in this part of the study: white noise, low-pass noise, high-pass noise, male-spectrum and female-spectrum shaped noise. In all cases the noise was recorded on Channel 2 of the tape recorders and the speech on Channel 1.

The white noise was obtained from a Grason-Stadler Model 901B noise generator. The spectrum of the recorded noise was checked using a 1/3 octave band analyzer. The results of this analysis are presented in Figure 2.

The high-pass noise was obtained by passing the output of the noise generator through an Allison AL-2ABR variable bandpass filter with a lower cut-off frequency of 500 Hz and an upper cut-off frequency of 8000 Hz. The rate of cut-off was 36 dB/octave. The upper cut-off frequency was chosen because of the limited response of tne TDH-49 earphone which would be used by the listeners. Figure 3 shows the earphone response as measured on a standard 6cc coupler. The frequency response of the filter system is shown in Figure 4 and the results of a 1/3 octave band analysis on the noise is also shown in Figure 2.

The low-pass noise was obtained using the same bandpass filter but setting a lower cut-off frequency at 50 Hz and the upper cut-off frequency at 500 Hz. See Figures 2 and 4 for the frequency response and 1/3 octave band analysis of the low-pass noise.

The female-spectrum shaped noise was obtained by shaping the spectrum of the noise to match as closely as possible the averaged female speech spectrum as measured for the five female talkers in the study. This was done by first obtaining the frequency power spectrum of each female talker. The CHABA recordings were subjected to a

Fig. 2. 1/3 octave band analysis of low-pass noise, high-pass noise and white noise.

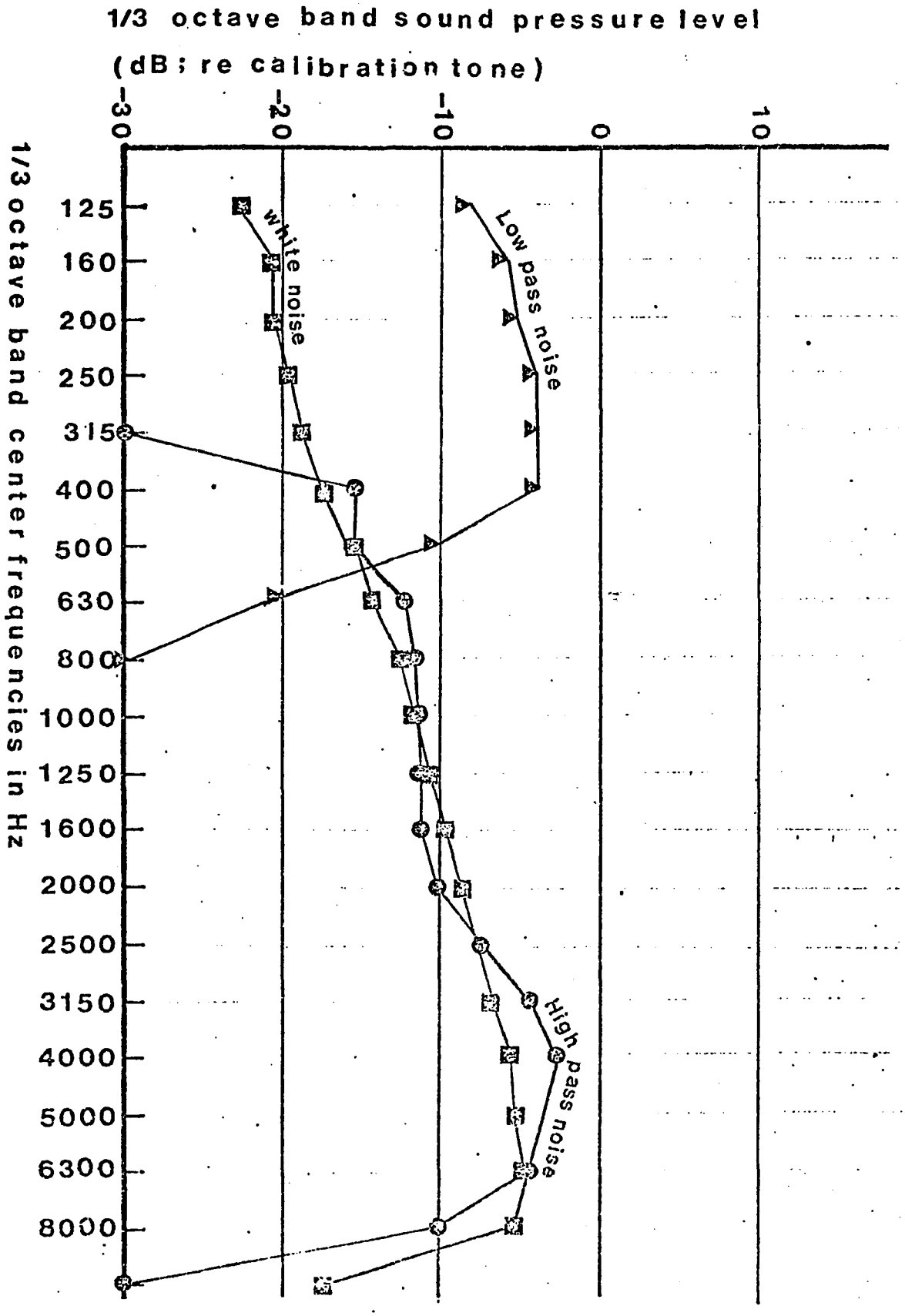

1/3 octave band sound pressure level
(dB; re calibration tone)
-30
-20
-10
0
10
1/3 octave band center frequencies in Hz
125
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
White noise
Low pass noise
High pass noise

Fig. 3. Graphic level recording of TDH-49 earphone response for earphone used in study.

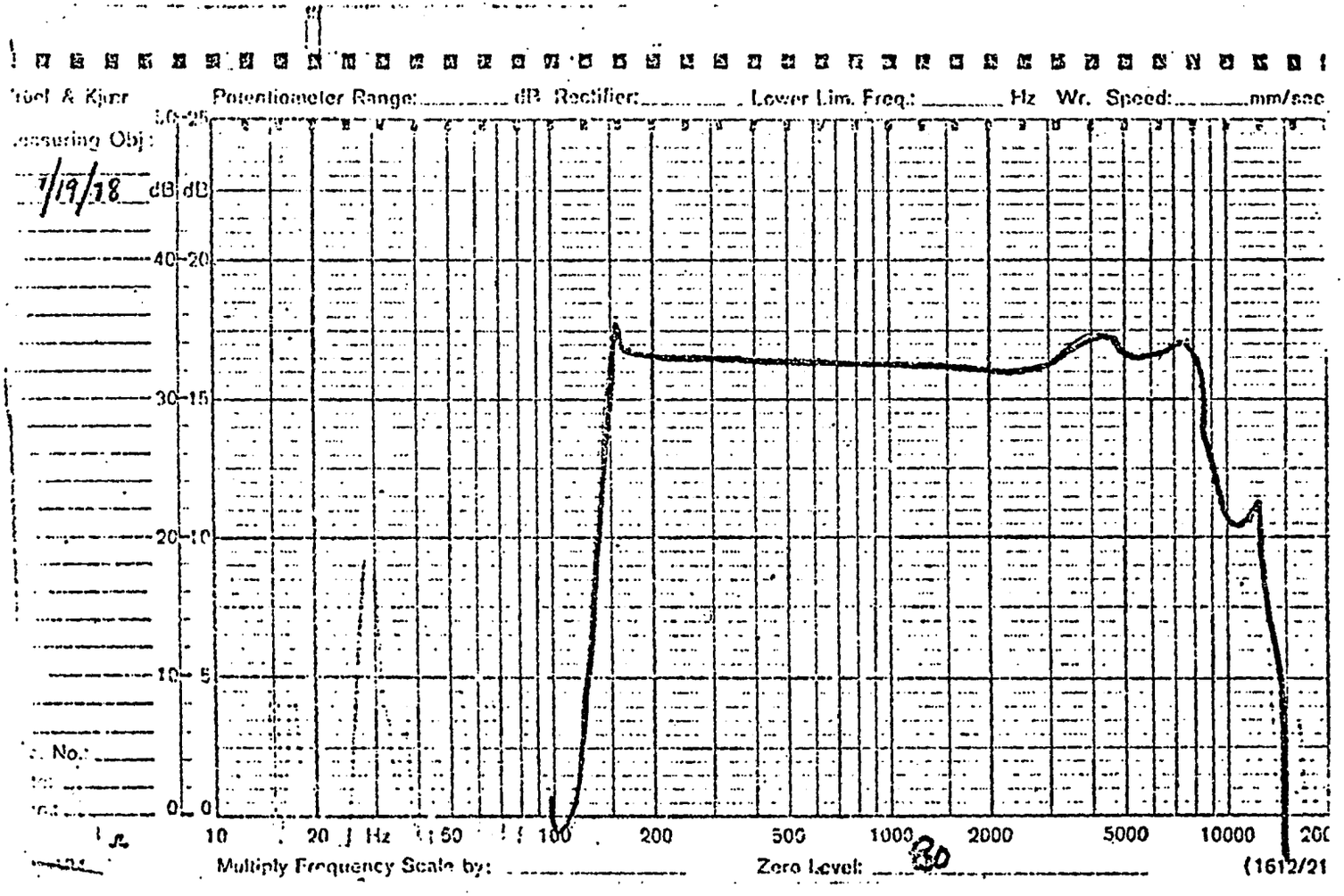
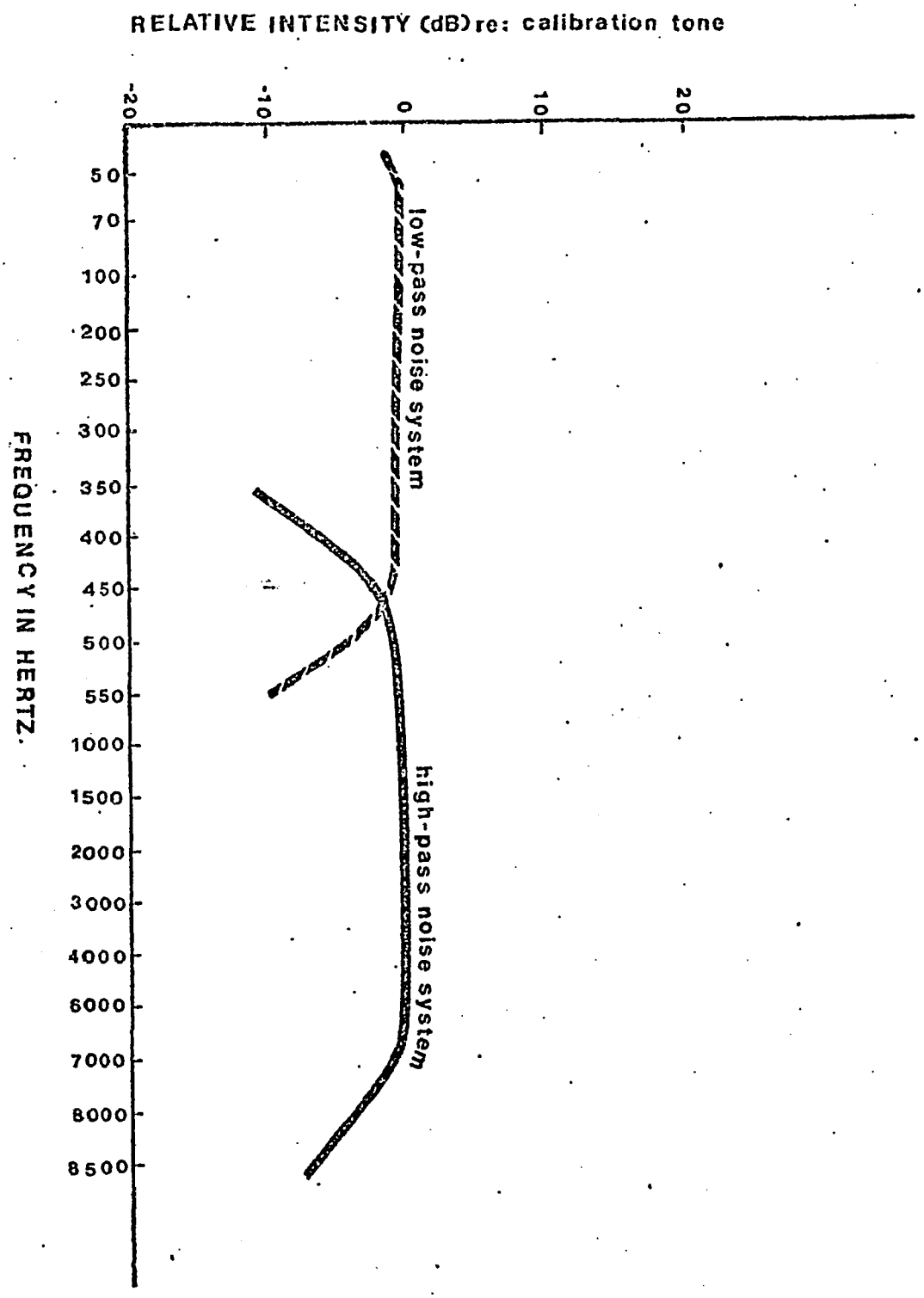

Potentiometer Range: dB Rectifier: Lower Lim. Freq.: Hz Wr. Speed: mm/sec
7/19/18
dB dB
40 20
30 15
20 10
10 5
0 0
10 20 Hz 50 100 200 500 1000 2000 5000 10000
Multiply Frequency Scale by:
Zero Level: 30

Fig. 4. Frequency response of the filter system used to obtain the high-pass and low-pass noise signals.


RELATIVE INTENSITY (dB) re: calibration tone
-20
-10
0
10
20
50
70
100
200
250
300
350
400
450
500
550
1000
1500
2000
3000
4000
6000
7000
8000
8500
low-pass noise system
high-pass noise system
FREQUENCY IN HERTZ.

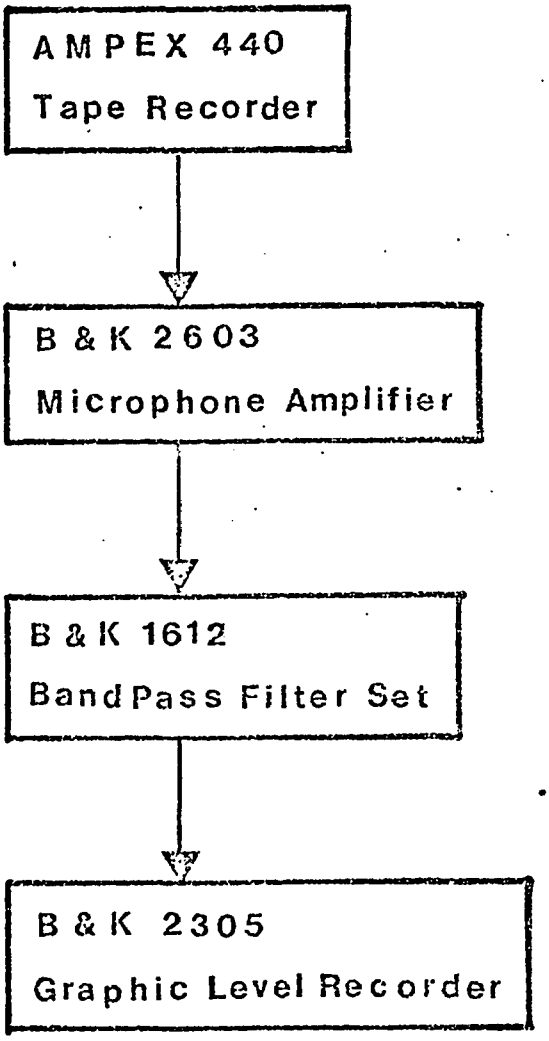
1/3 octave band analysis. The output of each 1/3 octave band was measured on a graphic level recorder. Figure 5 shows a block diagram of the analyzing equipment. The settings used on the microphone amplifier and graphic level recorder are shown in Appendix C. These settings were chosen such that the graphic level recorder would bracket the peak speech levels reliably with minimal overshoot. Measurements were not obtained for 1/3 octave bands above 8500 Hz because of the frequency response limitation imposed by the earphone.

The procedure used in making the graphic level recordings was as follows:

1. The input attenuator was set such that the 1000 Hz calibration tone at the beginning of each tape recording was set to produce a reading of 20 dB on the graphic level recorder. In doing this, the 1/3 octave band filter was also set to 1000 Hz.

2. The 1/3 octave band was set to the frequency band to be analyzed. The input attenuator of the graphic level recorder was then adjusted such that the speech peaks read between 10 and 100 dB on the graphic level recorder. The settings of the input attenuator for each graphic level recording were noted. Also identified on each recording was the talker and the 1/3 octave band setting. Typically two sentences per 1/3 octave band were recorded, but if an extremely short (i.e., two or three word) sentence occurred, then a third sentence was recorded.

Fig. 5. Block diagram of equipment used in the analysis of speech materials.


AMPEX 440
Tape Recorder
B & K 2603
Microphone Amplifier
B & K 1612
Band Pass Filter Set
B & K 2305
Graphic Level Recorder

The rule for measuring the tracings is shown in Figure 6. Only peaks within 20 dB of the highest peak were averaged on the tracing. A line is drawn on the tracing at the point 20 dB below the highest peak and the various peaks within this range were measured.

Figure 7 shows the speech spectra obtained for the five female talkers in the study. The average spectrum for all five female talkers was then obtained. This average spectrum is also shown in Figure 7.

The female-spectrum shaped noise was then produced by passing white noise from the Grason-Stadler 901B noise generator into a General Radio 1925 Multi-Filter in which the attenuator setting for each 1/3 octave band was adjusted so as to produce the spectrum shown in Figure 7. The frequency-shaped noise was then recorded on an Ampex 440 tape recorder. The recorded noise was finally subjected to a 1/3 octave band graphic level analysis. If the measured noise spectrum at any 1/3 octave band was not within $\pm$ 2 dB of the average female speech spectrum, the multifilter was adjusted and a new recording made. Figure 8 shows the average 1/3 octave band spectrum for the female talkers and the female-spectrum shaped noise. The two spectra do not differ by more than 2 dB at any 1/3 octave band.

The male-spectrum shaped noise was obtained using essentially the same procedure. Figure 9 shows the five male talker

Fig. 6. Illustration of method used to measure graphic level recordings of speech materials.

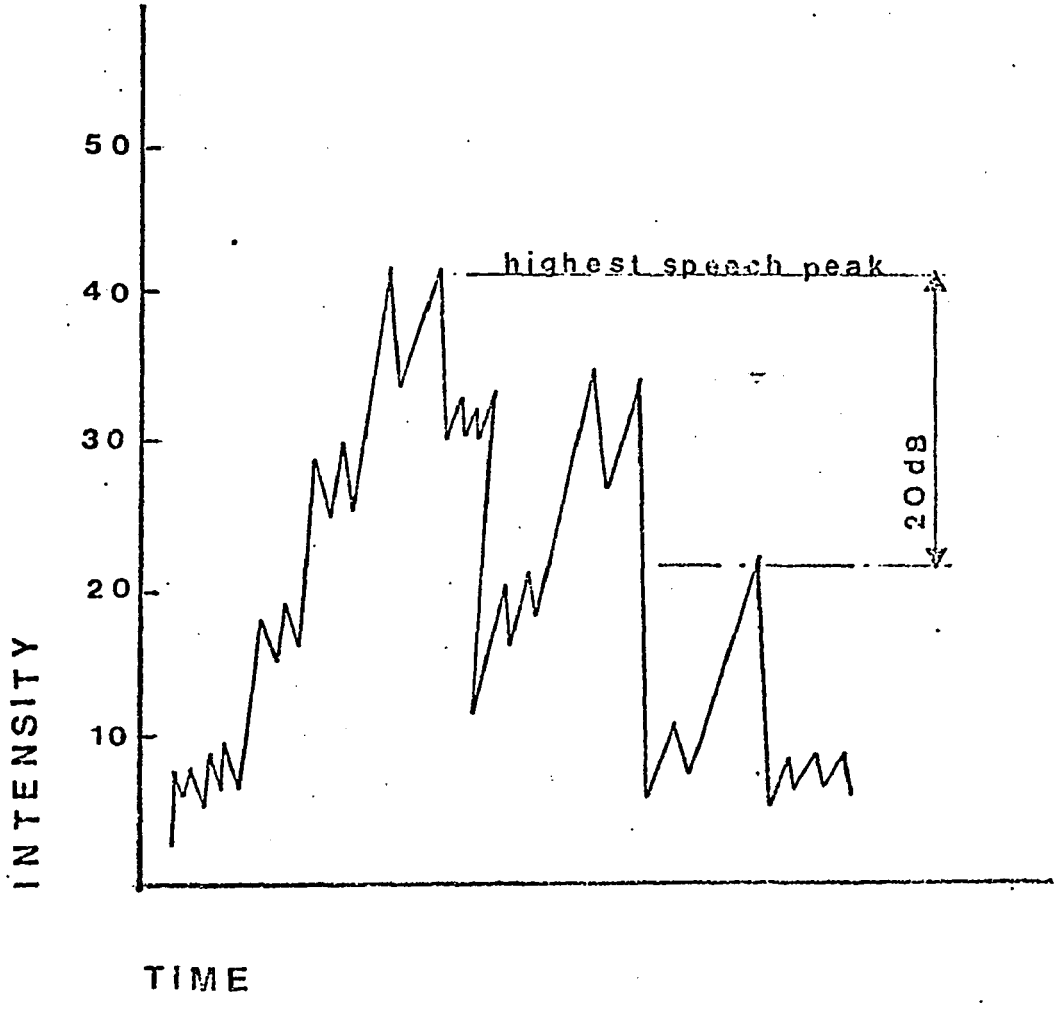

50
40
30
20
10
INTENSITY
highest speech peak
20dB
TIME

Fig. 7. 1/3 octave band speech spectra and average spectrum for the five female talkers in the study.

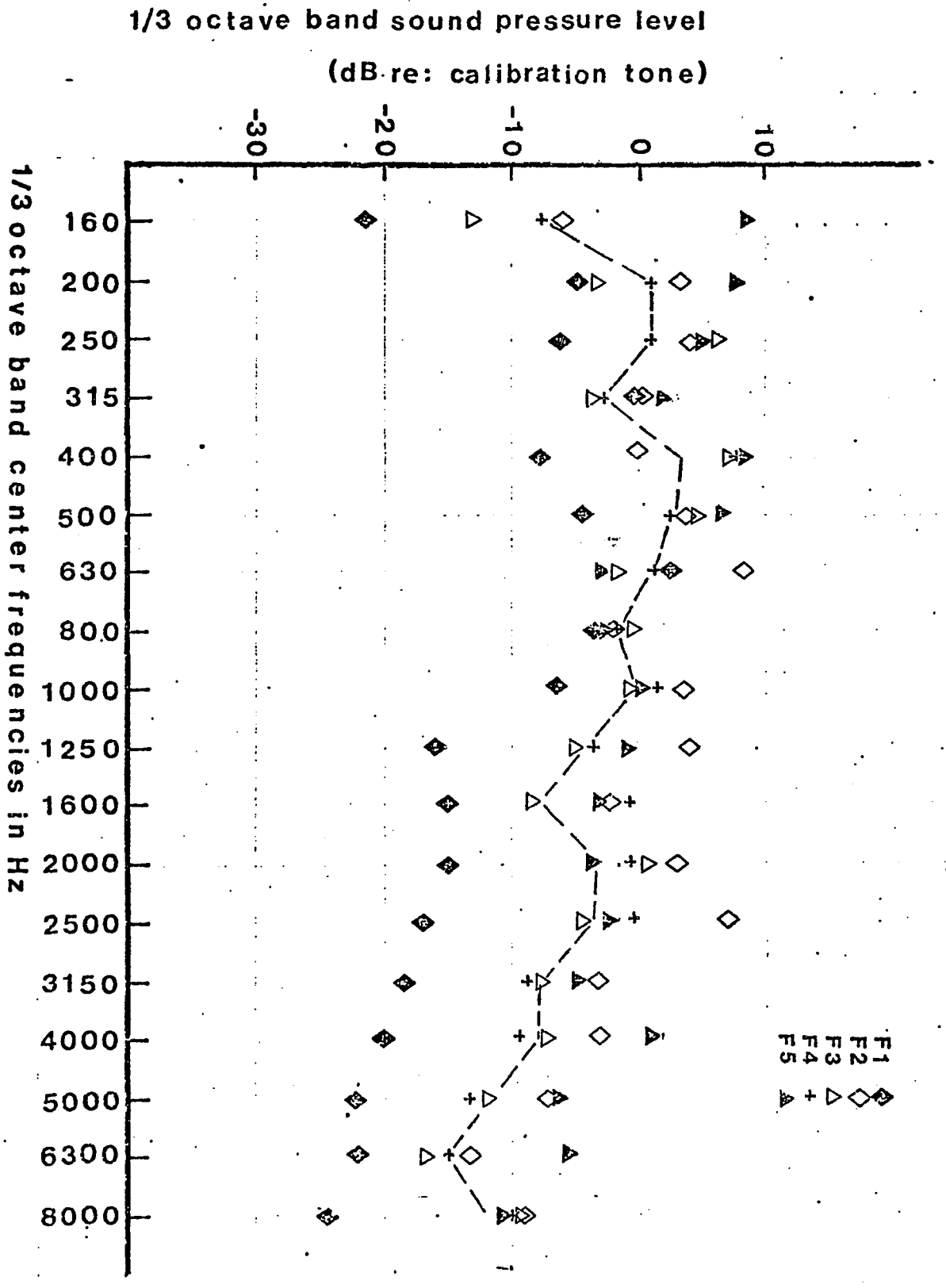

1/3 octave band sound pressure level
(dB re: calibration tone)
-30
-20
-10
0
10
1/3 octave band center frequencies in Hz
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
F1
F2
F3
F4
F5

Fig. 8. Average 1/3 octave band speech spectrum for the female talkers and the 1/3 octave band spectrum for the female-spectrum shaped noise.

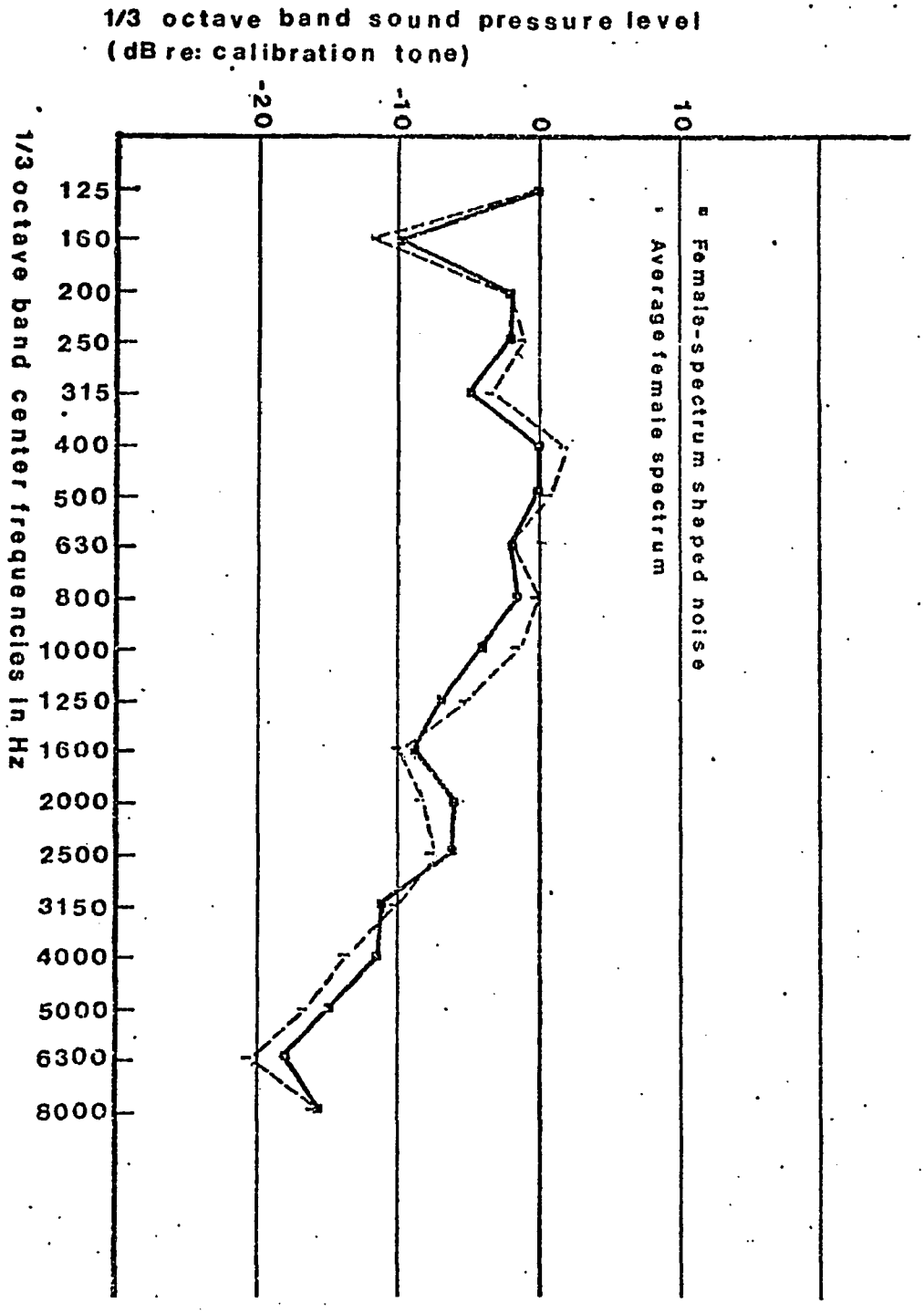
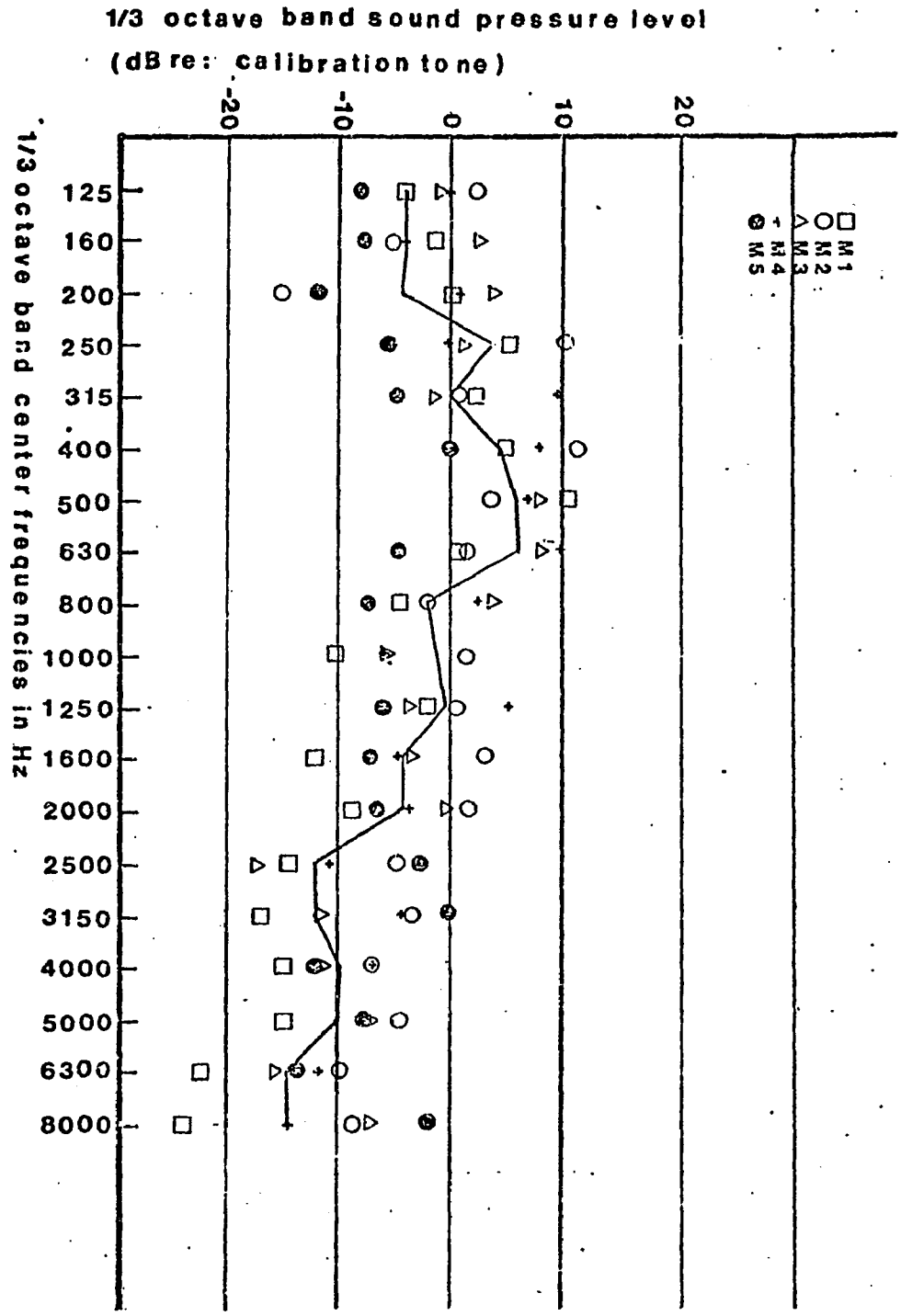

1/3 octave band sound pressure level
(dB re: calibration tone)
-20
-10
0
10
1/3 octave band center frequencies in Hz
125
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
Female-spectrum shaped noise
Average female spectrum

Fig. 9. 1/3 octave band speech spectra and average spectrum for the five male talkers in the study.


1/3 octave band sound pressure level
(dB re: calibration tone)
-20
-10
0
10
20
1/3 octave band center frequencies in Hz
125
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
M1
M2
M3
M4
M5

spectra and the averaged curve. Figure 10 shows the average 1/3 octave band spectrum for the male talkers and the male-spectrum shaped noise. Again it will be noted that the curves do not differ by more than 2 dB at any frequency.

Nonsense syllable materials

Eight randomizations of the nonsense syllable test were used in the study. Four of the versions were used with a female talker and four with a male talker. A different protocol was used for each of the four experimental conditions for each talker, as shown in Table 1. The listing of each of the experimental sequences is shown in Appendix D.

A total of eight nonsense-syllable test tapes were prepared, a male and a female version of the conditions listed in Table 1. The preparation of the master tapes is detailed below:

0 dB/octave slope condition. Recordings of the NST protocol #8 were made for both the male and female talker.[2] A 1000 Hz calibration tone was recorded at the beginning of each tape.

0 dB/octave slope +speech spectrum noise condition. Recordings of Protocol #6 male and female were used for the two

[2]The protocol numbers listed here are those assigned to the eight protocols developed and used by Resnick et al. (1975).

Fig. 10. Average 1/3 octave band speech spectrum for the male talkers and the 1/3 octave band spectrum for the male-spectrum shaped noise.

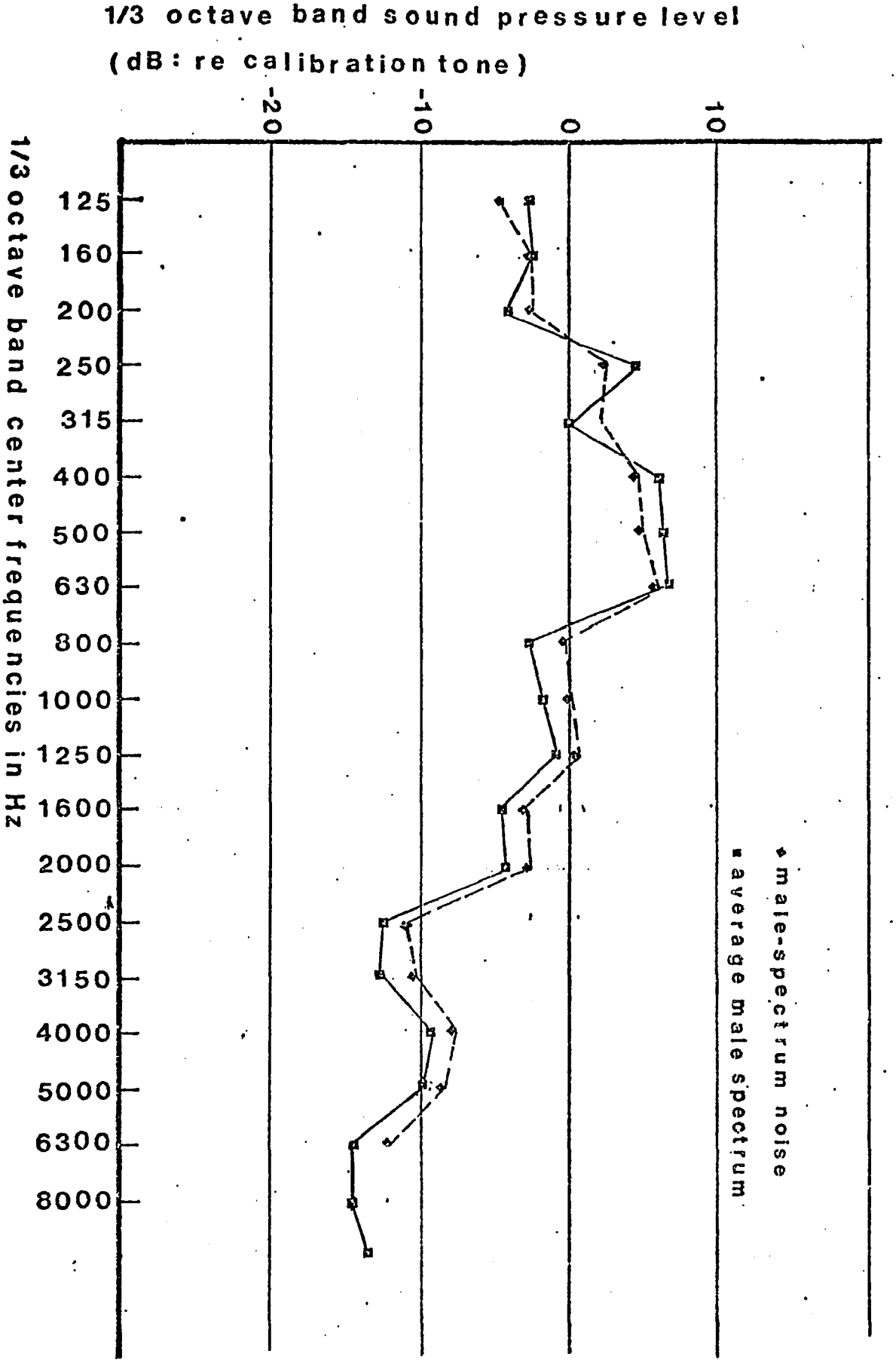

1/3 octave band sound pressure level
(dB: re calibration tone)
-20
-10
0
10
1/3 octave band center frequencies in Hz
125
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
♦ male-spectrum noise
▪ average male spectrum

TABLE 1

PROTOCOL ASSIGNMENTS FOR THE NONSENSE SYLLABLE TEST FOR THE DIFFERENT EXPERIMENTAL CONDITIONS

Protocol*	Talker	Experimental Condition
5	Male	+6 dB/octave slope
5	Female	+6 dB/octave slope
6	Male	0 dB/octave slope + noise
6	Female	0 dB/octave slope + noise
7	Male	-6 dB/octave slope
7	Female	-6 dB/octave slope
8	Male	0 dB/octave slope
8	Female	0 dB/octave slope

*The protocol numbers listed here are those assigned to the eight protocols developed and used by Resnick et al. (1975).

noise conditions. Female-spectrum shaped noise was recorded on the second track of the tape for the female talker. Similarly, male-spectrum shaped noise was recorded on the second track for the male talker.

+6 dB/octave slope condition. Recordings of Protocol #5 male and female were used for this condition. The recordings were made using the equipment shown in the block diagram of Figure 11. An 8000 Hz cut-off was used to limit the speech bandwidth to be within the frequency bandwidth of the headphone. The frequency response curve of the system is shown in Figure 12.

-6 dB/octave slope condition. Recordings of Protocol #7 male and female were used for the -6 dB/octave slope condition. The recording procedure was the same as that used for the +6 dB/octave slope condition, except that it was unnecessary to include the stereo amplifier into the system because the signal intensity was sufficiently high without amplification. The frequency response curve for this condition is also shown in Figure 12. Copies of these eight tapes were used for data collection.

Preparation of Calibration Materials

The calibration tapes prepared for this study served two distinct functions. First, they were used to verify that the flat, upward

Fig. 11. Block diagram of equipment used to produce the filtered nonsense syllable materials.

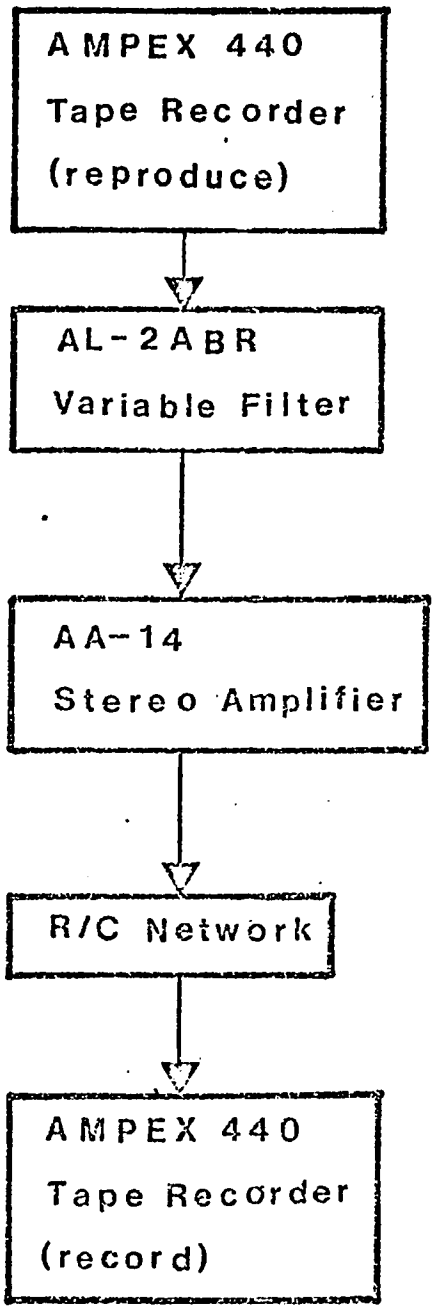
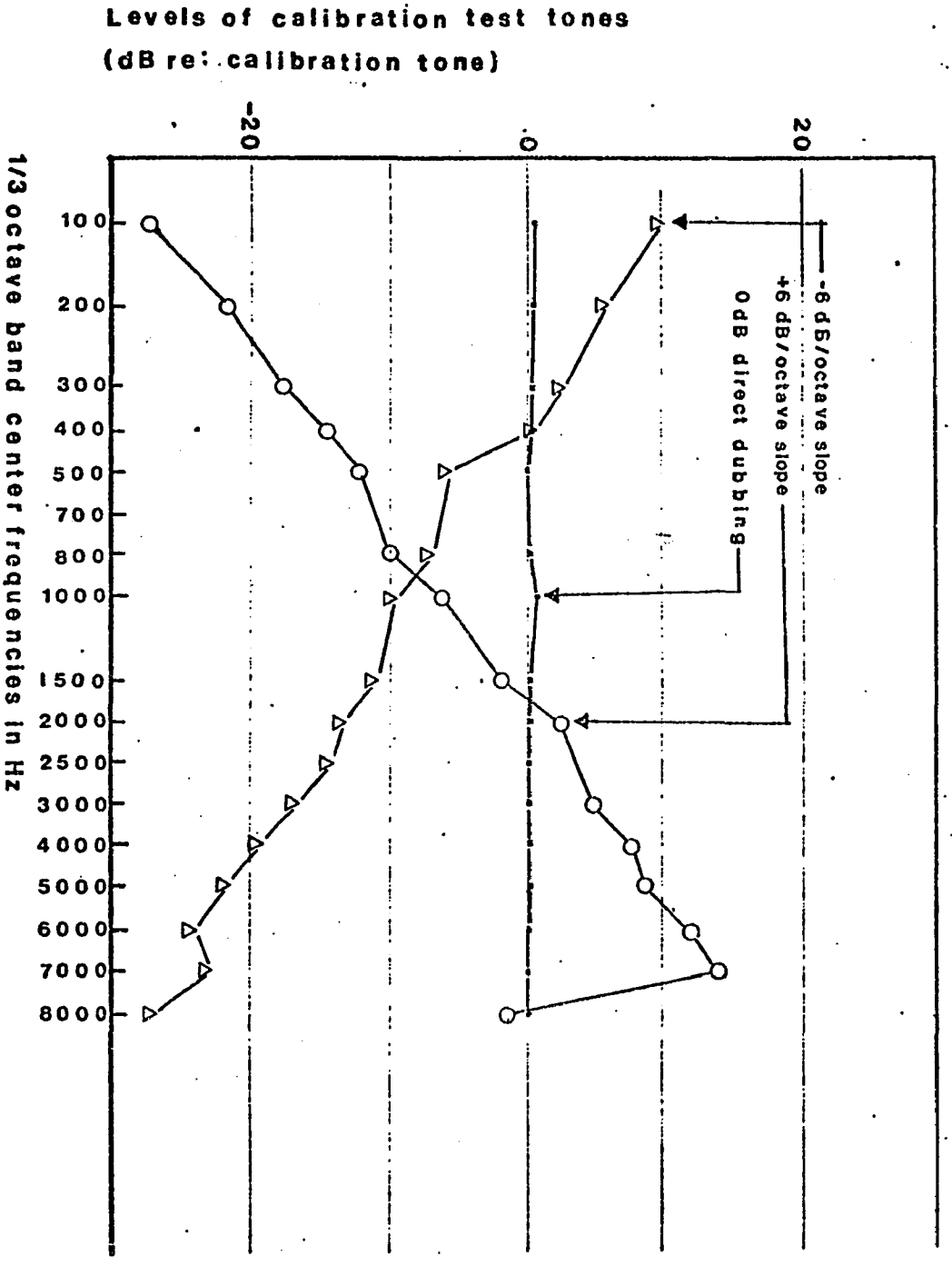

AMPEX 440
Tape Recorder
(reproduce)
AL-2ABR
Variable Filter
AA-14
Stereo Amplifier
R/C Network
AMPEX 440
Tape Recorder
(record)

Fig. 12. Frequency responses of the filter systems used to obtain the filtered nonsense syllable materials (-6 dB/octave slope and +6 dB/octave filter slope).


Levels of calibration test tones
(dB re: calibration tone)
-20
0
20
1/3 octave band center frequencies in Hz
100
200
300
400
500
700
800
1000
1500
2000
2500
3000
4000
5000
6000
7000
8000
-6 dB/octave slope
+6 dB/octave slope
0 dB direct dubbing

and downward sloping frequency responses used for the nonsense syllable conditions were processing the speech appropriately. In addition, a calibration tape was also used to check the frequency response characteristics of the recording equipment at the laboratory, the playback equipment at the laboratory, and the playback equipment at the hospital site.

In all, three different calibration tapes were prepared, each corresponding to one of the three filter conditions used for the nonsense syllable materials. Each of these tapes consisted of tonal signals twenty seconds in duration at frequencies of 100, 200, 300, 400, 700, 1000, 1500, 2000, 2500, 3000, 4000, 5000, 6000, 7000, and 8000 Hz recorded at a constant input voltage.

In each case the preparation of the calibration tape recreated the procedure used in preparing the corresponding speech test tape. The only difference in the procedure was the substitution, at the recorder input, of a sine wave signal in place of the speech signal. Figure 12 shows the results of a 1/3 octave band analysis of the calibration tapes. This analysis is the same as that used for checking both the spectra of all speech and noise recordings. The analysis shown in Figure 12 depicts the frequency response characteristic of the overall system from playback recorder to the final reproduction stage. In addition to checking the recording conditions, the equipment at the hospital site was checked using the 0 dB/octave

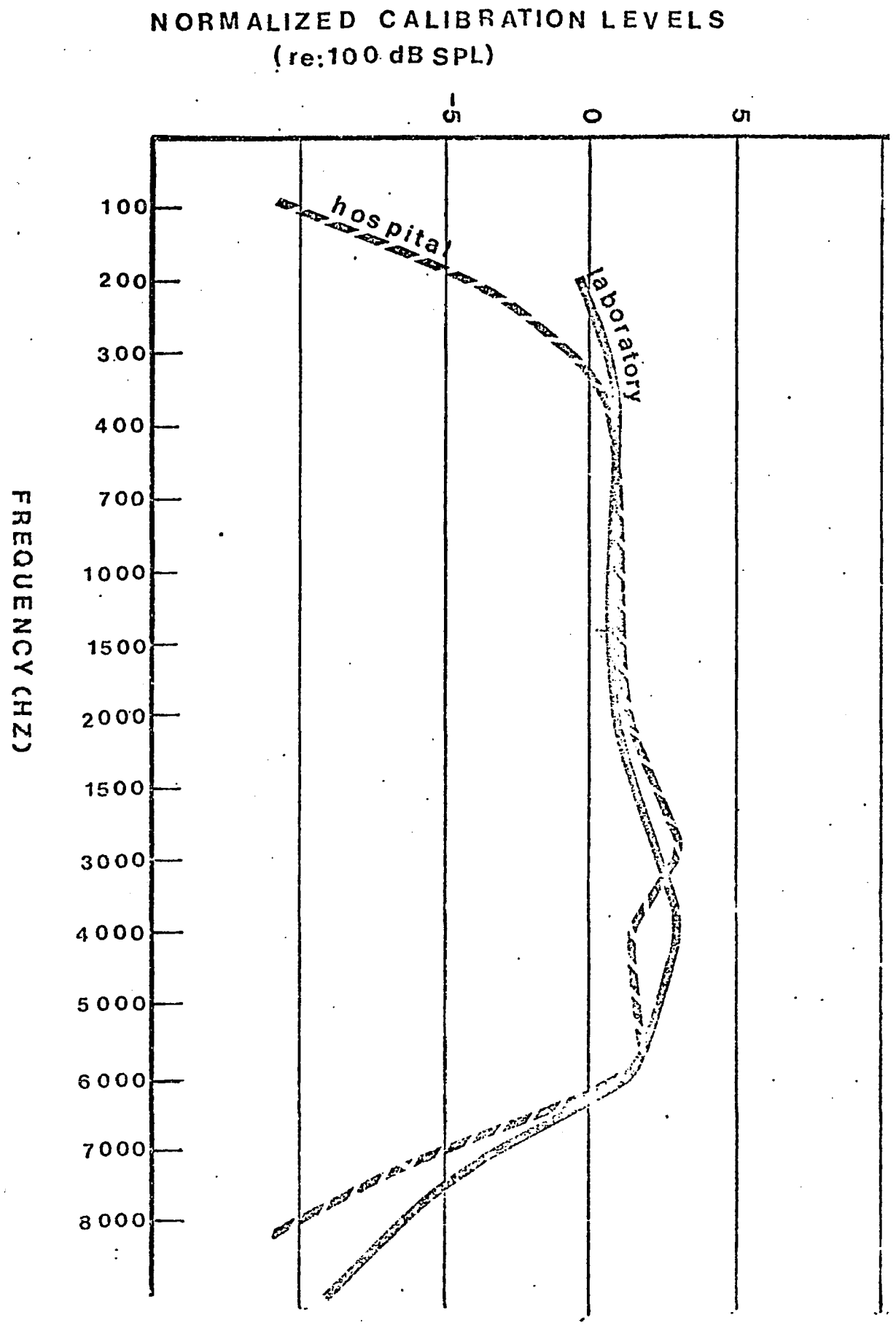
slope calibration tape.

The on-site equipment was calibrated the day before the data collection was scheduled to begin. Figure 13 presents the results of the two sets of measurements obtained with this tape, at the laboratory and at the hospital location. The differences in frequency response between these two curves are within $\pm$ 2 dB over the frequency range 100-7500 Hz. This was considered adequate for the purposes of this study.

Test procedure

Each subject was tested individually. The subject was seated in the test room of an IAC 400 Series two room audiometric test suite which met ANSI-1969 specifications for ambient noise level. All testing was completed in a single two and one half hour session. The initial portion of the session was reserved for the basic audiological evaluation. This included pure-tone air and bone conduction threshold measurements (250-8000 Hz) as well as SRT, speech discrimination measures and estimates of the subject's most comfortable listening level and the threshold of discomfort. These data were obtained using the Grason-Stadler 1701 audiometer and SONY TC-377 tape recorder with TDH-49 earphones in MX-41AR cushions. A tympanogram was obtained on each ear using an American Electromedics acoustic impedance bridge Model 83. A copy of the record sheet used for the

Fig. 13. Results of calibration tape comparison measurements obtained at the laboratory and hospital sites.


NORMALIZED CALIBRATION LEVELS
(re:100 dB SPL)
-5
0
5
100
200
300
400
700
1000
1500
2000
1500
3000
4000
5000
6000
7000
8000
FREQUENCY (HZ)
hospital
laboratory

study is presented in Appendix E.

Any other information which was not readily available from the subject's departmental chart was also obtained at this time (i.e., date of birth, surgical history, hearing aid experience, etc.).

At the conclusion of this initial evaluation, a decision was made whether or not the subject met the criteria for selection. If so, the better ear was selected as the test ear for subsequent testing.

Materials were presented to the subjects at MCL. The method of obtaining MCL is discussed in the next section. A -15 dB S/N ratio was used for the normal hearing and conductive hearing loss groups for the CHABA sentences and a 0 dB S/N ratio was used for the sensori-neural hearing loss groups. A 0 dB S/N ratio was used for the noise conditions of the nonsense syllable test for all listeners. A complete listing of the MCL values for all conditions for all subjects is presented in Appendix F.

Half-way through the nonsense syllable runs a 10-15 minute break was offered to all subjects to avoid fatigue and increase concentration on the listening task.

Method of obtaining MCL

Initially, an approximation was made as to the comfortable speech level which the subject should be able to hear on the basis of the audiogram and SRT. This was necessary because it was found,

especially for the severely sloping sensori-neural group, that the SRT alone was not a sufficient indicator of MCL. Using this level, the listener was queried as to whether the voice of the examiner was audible. Once this was determined, the subject was given the following instructions:

> You are going to hear a male or female voice repeating either "You will mark" or continuous material; you are to indicate when the speech is at a level which is most comfortable for you to listen. You do not need to repeat any of the things you hear; just indicate when the voice is at a level at which you feel comfortable. You may indicate by raising your hand or telling me.

The attenuator was then slowly increased until the subject signalled that the level was comfortable. In most cases the level was increased 5-10 decibels and the listener was asked about that level for comfort, just as a means to verify that the comfort level had not been prematurely indicated. The study subjects were generally consistent in their comfort level choices and where previous test results were available, the results were essentially similar (i.e., within 5-10 dB between previously and presently obtained MCL's).

Once the MCL was determined for the particular experimental condition, the subject was given the instructions for that task.

For the nonsense syllable tests the subject was given the test

booklets in the protocol presentation order. The booklet for the first list was handed to the subject and it was explained that he would be listening to speech sounds both in quiet and in noise. While some of the syllables might sound like words, they were not, and were to be thought of as nonsense syllables.

They were shown that all possible answers for each test item were on the page. They were instructed to circle or mark the syllable that they heard in each instance and to mark only one syllable for each item.

It was explained that they would hear either a male voice or a female voice which would indicate the page in the test booklet. Then the voice would say "You will mark _____, please." The test syllable would follow the word "mark" and they were to indicate on the test page the syllable they had heard. Upon completion of the items on Page 1, they would turn to Page 2 and proceed in the same manner until the booklet was completed. A sample sheet is provided in Appendix G.

It was stressed that they were to mark exactly what they heard even if in some cases they had heard the same sound more than once and in other cases to guess if they had not heard a particular sound at all.

In those nonsense syllable conditions where noise was introduced, the subjects were additionally instructed not to be concerned if

the task seemed unusually difficult. It was explained that we knew the task was difficult and that if they became confused or lost their place on the page to so indicate and the test tape would be stopped.

For the CHABA sentences, the subjects were informed that they were going to hear a number of sentences spoken by different talkers, all in the presence of noise. At the completion of each sentence they were requested to repeat exactly what they heard. As many of the words were to be repeated as possible, whether or not the resulting sentence made any sense. It was stressed that even if only one word was heard, it was to be repeated. The subjects' responses to each sentence were recorded by the experimenter. A sample score sheet will be found in Appendix H.

Following each CHABA sentence list or nonsense syllable run, the subject was asked to rate the intelligibility of that talker. He was asked to do so on a scale of 1-7. A rating of 1 meant that the talker was poor with respect to intelligibility and a rating of 7 meant that the talker was highly intelligible. This rating procedure was done for each of the ten CHABA lists and the eight nonsense syllable variations. Rating sheets for both the sentence and nonsense syllable materials will be found in Appendix I.

Human subjects protection

Hospital clients who were considered potential study subjects

were initially contacted by an introductory letter requesting their participation in the study. This letter was followed by a telephone call from this researcher. Those individuals who were not interested in participating in the study were not contacted again. Those who were willing to participate were given specific appointment times for the testing. At the time of this appointment, the study was explained to the individual and a signed consent form was obtained. A thank you letter was sent to all participants upon completion of the study. Copies of the letter sent to the subjects, the consent form and the thank you letter are included in Appendix J.

CHAPTER IV

RESULTS

Preliminary Analysis

The study involved two different types of speech reception materials, sentences and nonsense syllbables. Different experimental conditions were used with each type of material. The sentences were always presented in the presence of one of five different types of noise, and the nonsense syllables were presented in four different ways: three frequency response conditions in quiet and one noise condition. The five listener groups were subdivided into older and younger listeners (five oldest vs. five youngest in each group with the median age approximately fifty years old), as well as into male and female listeners. The factors that were considered in the analysis of the sentence data were: Noise Type, Talker Gender, Age, Experimental Group and Listener Gender. The factors considered for the nonsense syllable data were: Experimental Condition, Talker Gender, Age, Experimental Group and Listener Gender.

After all the data were analyzed it was discovered that the sentence recordings for one of the five male talkers was produced at a slightly lower signal-to-noise ratio (-5 dB) than the other recordings

which were at 0 dB. In order to determine whether this disparity would affect the major results, the scores for this talker and a corresponding female talker who did the same experimental conditions were eliminated from the data and the analyses of variance repeated using mean scores across conditions for the varying experimental groups. It was found that while there were small differences between the mean scores for the two analyses (on the order of 2-3 percentage points as shown in Table 2), the main effects and statistically significant trends were the same whether four or five talkers were used in the analysis.

It was desirable to use the five talker data, however, since the male talker with the lower signal-to-noise ratio was the same talker for whom nonsense syllable data were available for the subjects in the study. In order to correct for the 5 dB difference on the sentence materials, additional data were collected for a new group of subjects in which both the 0 dB and -5 dB signal-to-noise ratios were used on CHABA sentences for this talker. Five listener groups of five subjects who matched the experimental criteria were used for this supplementary study.

From these data the effect of a 5 dB change in signal-to-noise ratio was determined for the same noise, experimental group, and talker. The data of the main study for this talker under these conditions was then adjusted for these changes. The original scores

TABLE 2

MEAN PERCENTAGE SCORES FOR COMPARISON OF SENTENCE DATA FOR FOUR TALKERS PER GENDER GROUP VS. FIVE TALKERS PER GENDER GROUP

		4 Talkers	5 Talkers
Variable	Condition	% Correct	% Correct
Gender	Male	54	53
	Female	63	67
Experimental Condition	High-pass	64	64
	Low-pass	86	86
	White	68	71
	Male-spectrum	35	39
	Female-spectrum	39	42
Experimental Group	Conductive H.L.	56	57
	Normal hearing	49	51
	Flat sensori-neural	65	68
	Moderately sloping sensori-neural	64	65
	Severely sloping sensori-neural	59	62

as well as the corrected scores are presented in Appendix K.

As a check, the data were analyzed both for the original data (with the lower signal-to-noise ratio) as well as for the adjusted data. The two analyses provided essentially the same results, in that the same main effects and second-order interactions were found to be statistically significant. The magnitude of the difference varied by approximately 5 percent; the major difference between the two analyses was that there were more statistically significant higher-order interactions for the adjusted data. From this point on, only the sentence data with the adjustments are considered. The syllable data did not involve any adjustments.

Main Analysis

Effect of Listener Gender

As was noted earlier it was of interest to test for Listener Gender. Because of limitations in the available subject pool, it was not possible to have exactly five subjects of each gender in each of the cells of the experimental design. The number of subjects in each cell varied from a low of three to a high of eight, although in the majority of cases there were five listeners per cell. However, if one pooled the data over the Age factor, one could test for Listener Gender in a balanced design of ten subjects per cell thereby allowing for a more powerful statistical test for the effect of Listener Gender in all

combinations of variables except Age. This was done recognizing that the analysis might not detect any interaction between Age and Listener Gender. This was done in a four factor analysis of variance where the factors considered were Listener Gender, Talker Gender, Experimental Condition, and Experimental Group. The scores for different listeners were treated as replications.

The results showed that for neither the sentences nor the nonsense syllables was Listener Gender a significant variable. It was therefore decided that the male and female listeners could be grouped together for the remainder of the analyses. This allowed for more powerful statistical tests for all of the remaining variables. It should also be noted that since no significant effects were obtained for Listener Gender (neither main effects nor any interactions with other variables), it seemed unlikely that there might be an interaction between Listener Gender and Age--the one interaction that could not be tested directly in this analysis.

Sentence intelligibility: Subjects with hearing loss

An analysis of variance was performed on the sentence scores for all hearing-impaired subjects. Four factors were considered: Type of Noise (high-pass, low-pass, white, male-spectrum and female-spectrum shaped noise), Talker Gender (male or female), Age (five oldest vs. five youngest per group), Experimental Group

(conductive, flat sensori-neural, moderately sloping sensori-neural, and severely sloping sensori-neural hearing loss). There were ten subjects per experimental group.

The results of the analysis of variance are shown in Table 3. The arc-sine transformation was used in order to stabilize the error variance (Edwards, 1958). A key assumption of the analysis of variance is that the error variance is the same for every cell. It can be shown, however, that estimates of a proportion have a variance that is dependent on the value of the proportion. If the arc-sine transformation is used, the variance is approximately $1/n$ where n is the number of items in the test.

The age and hearing loss factors necessarily involved different subjects. Hence, in testing for statistical significance of these factors (and any interactions involving these factors), the between-subject variation was used as the residual variance (i.e., the denominator in the F-test). The Noise and Talker Gender factors (and Noise-Gender interaction) were varied over their entire range for each subject. In this case, the highest order interaction in the analysis of variance provided an estimate of the within-subject variability and was used as the denominator in the F-test. It is important to recognize that there are limitations to the degree that one can generalize about the observations relating to gender effects.

Since every subject received every talker under every

TABLE 3

ANALYSIS OF VARIANCE FOR SENTENCE SCORES OF HEARING-IMPAIRED LISTENERS

Source of Variation	Degrees of Freedom	Mean Squares	F-Ratio	Signif. Level
Noise (N)	4	.39	54.07	.001
Talker Gender (G)	1	.13	18.38	.001
NG Interaction	4	.01	2.63	.005
Within subject variation	48	.00		
Age (A)	1	.05	7.61	.001
Experimental group (P)	3	.04	5.91	.001
AP Interaction	3	.14	19.29	.001
NA Interaction	4	.00	1.18	.010
GA Interaction	1	.00	.61	
NGA Interaction	4	.00	1.34	
NP Interaction	12	.01	2.12	.005
GP Interaction	3	.00	.24	
NGP Interaction	12	.01	1.41	.001
NAP Interaction	12	.02	3.45	.001
GAP Interaction	3	.00	.15	
NGAP Interaction	12	.01	2.03	.005
Between subject variation	272	.01		

condition, statistically powerful tests are available for establishing average differences between the five male and five female talkers used in the study. It is much more difficult to establish whether male talkers in general are more intelligible than female talkers selected at random from the population because of the very large differences between talkers. This point will be discussed again later when the scores for the individual talkers are compared.

As shown in Table 3, all the main effects were found to be statistically significant. In addition, there were significant interactions between Noise-Talker Gender, Age-Experimental Group, Noise-Experimental Group, and Noise-Age. A three-way interaction was indicated between Noise-Age-Experimental Group and Noise-Gender-Experimental Group. There was one four-way interaction indicated between Noise-Gender-Age-Experimental Group. The effects found to be statistically significant are described below.

Noise (Experimental Condition). The data on effects of masking noise are summarized in Table 4. The mean scores for the five types of masking noise have been ranked in terms of effectiveness of masking, with the male and female-spectrum shaped noise being the most effective maskers and the low-pass noise providing the least effective masking.

Talker Gender. Significant talker gender differences were

TABLE 4

SENTENCE INTELLIGIBILITY SCORES FOR HEARING IMPAIRED LISTENERS IN THE PRESENCE OF MASKING NOISE

Masking Noise	Percent Correct
Male-spectrum shaped noise	44.3
Female-spectrum shaped noise	46.1
High-pass noise	63.7
White noise	75.4
Low-pass noise	81.8

obtained ($p < .001$), the male talkers obtaining a 57.0% mean score as compared with the 65.9% mean score for the female talkers.

Age. The effect of age was of smaller magnitude but not unimportant. It was found that the scores for the younger subjects were higher than for the older subjects (64.2% vs. 58.6%) when averaged over all experimental groups. As shall be seen later, however, this advantage for the younger listeners does not hold across all four hearing impaired groups of listeners.

Experimental Group. In comparing the data for the four groups of hearing impaired listeners, the conductive hearing loss group should be considered separately since a lower S/N ratio (-15 dB) was used for masking the sentence materials. As a result, the average scores for this group of listeners were well below those of the other listening groups, as shown in Table 5.

Age-Experimental Group Interaction. The advantage shown for the younger listeners does not hold across all four experimental groups. It was found that there was a reversal of this advantage for the conductive hearing loss group where the older listeners had an 18 percent higher mean score than the younger group. For the sensori-neural hearing loss groups the older listeners consistently scored lower than their younger counterparts with the differences

TABLE 5

SENTENCE INTELLIGIBILITY SCORES FOR FOUR EXPERIMENTAL GROUPS OF LISTENERS

Experimental Group	Percent Correct
Conductive hearing loss	55.3
Flat sensori-neural loss	67.5
Moderately sloping sensori-neural loss	62.3
Severely sloping sensori-neural loss	60.6

ranging from approximately 8 percent to 28 percent. Figure 14 presents the results of the Age-Experimental Group comparison. It should be noted that in this and all future figures (and tables) showing statistically significant interactions, the data are represented in terms of percent-correct score. The analysis in which the statistical significance was established, however, was done in terms of an arc-sine transformation of the data.

Noise-Gender Interaction. In none of the five different noise situations does the male talker obtain as high a mean score as the female talker. The differences in score, however, varied between noise conditions as is seen in Figure 15. The relative intelligibility of the female talkers compared to the male talkers improved systematically from male-spectrum shaped noise (less than 1%) to high-pass (4.6%), white (13%), female-spectrum shaped (13.4%) to low-pass (16.5%) noise. The numbers in parenthesis indicate the difference in percent score between the male and the female talkers.

Noise (Experimental Condition)-Experimental Group Interaction. Figure 16 shows the average scores for each Experimental Group for each Experimental Condition. On average, the highest scores were obtained for the low-pass noise, followed by the white and high-pass noise and then the male-spectrum and female-spectrum shaped noise. However, the ranking of scores by Experimental Group

Fig. 14. Mean percent scores for the age-experimental group interaction on the analysis of variance for the sentence data.

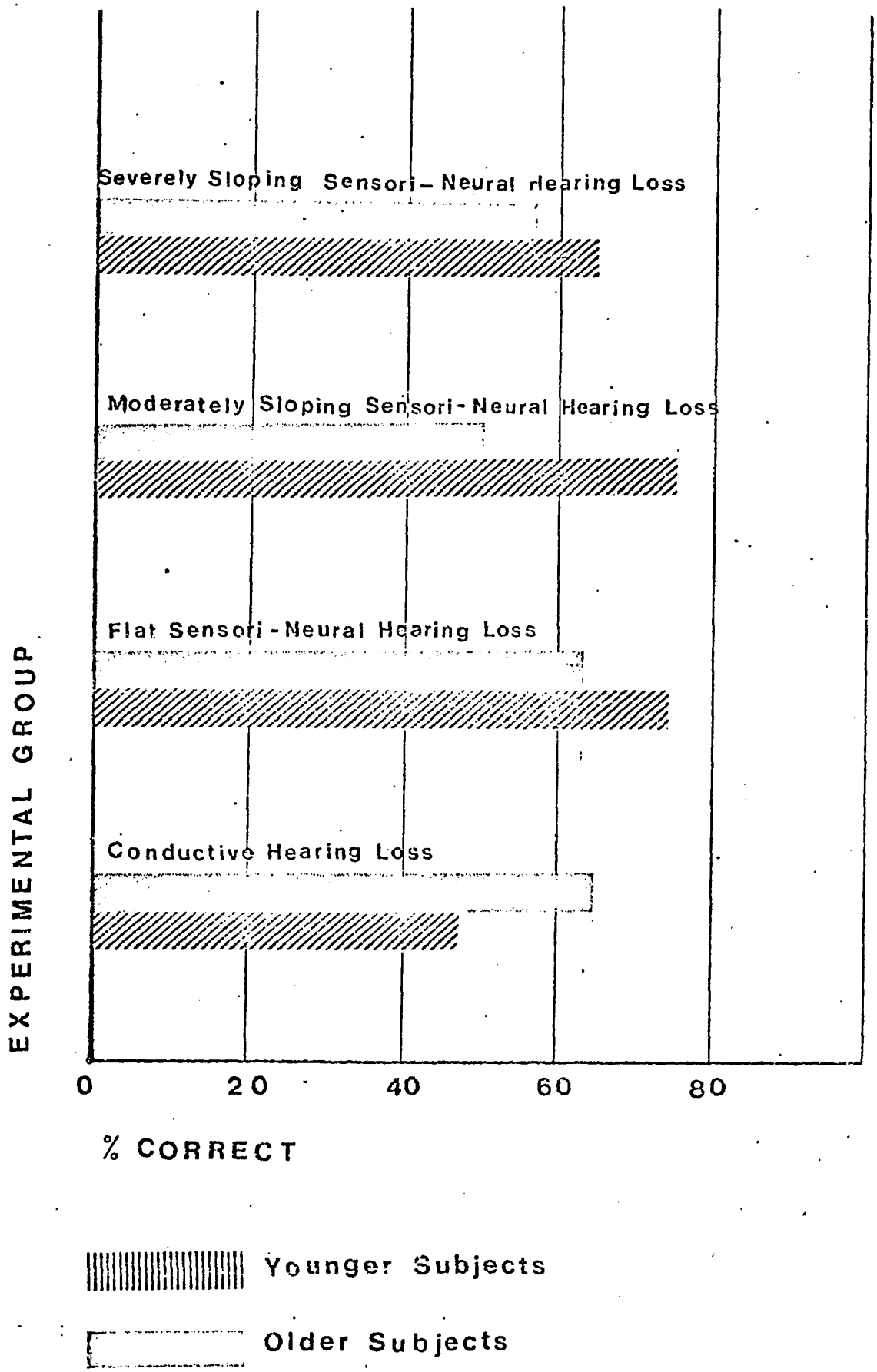

Severely Sloping Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Flat Sensori-Neural Hearing Loss
Conductive Hearing Loss
EXPERIMENTAL GROUP
0
20
40
60
80
% CORRECT
Younger Subjects
Older Subjects

Fig. 15. Mean percent scores for the experimental condition (noise)-gender interaction on the analysis of variance for the sentence data.

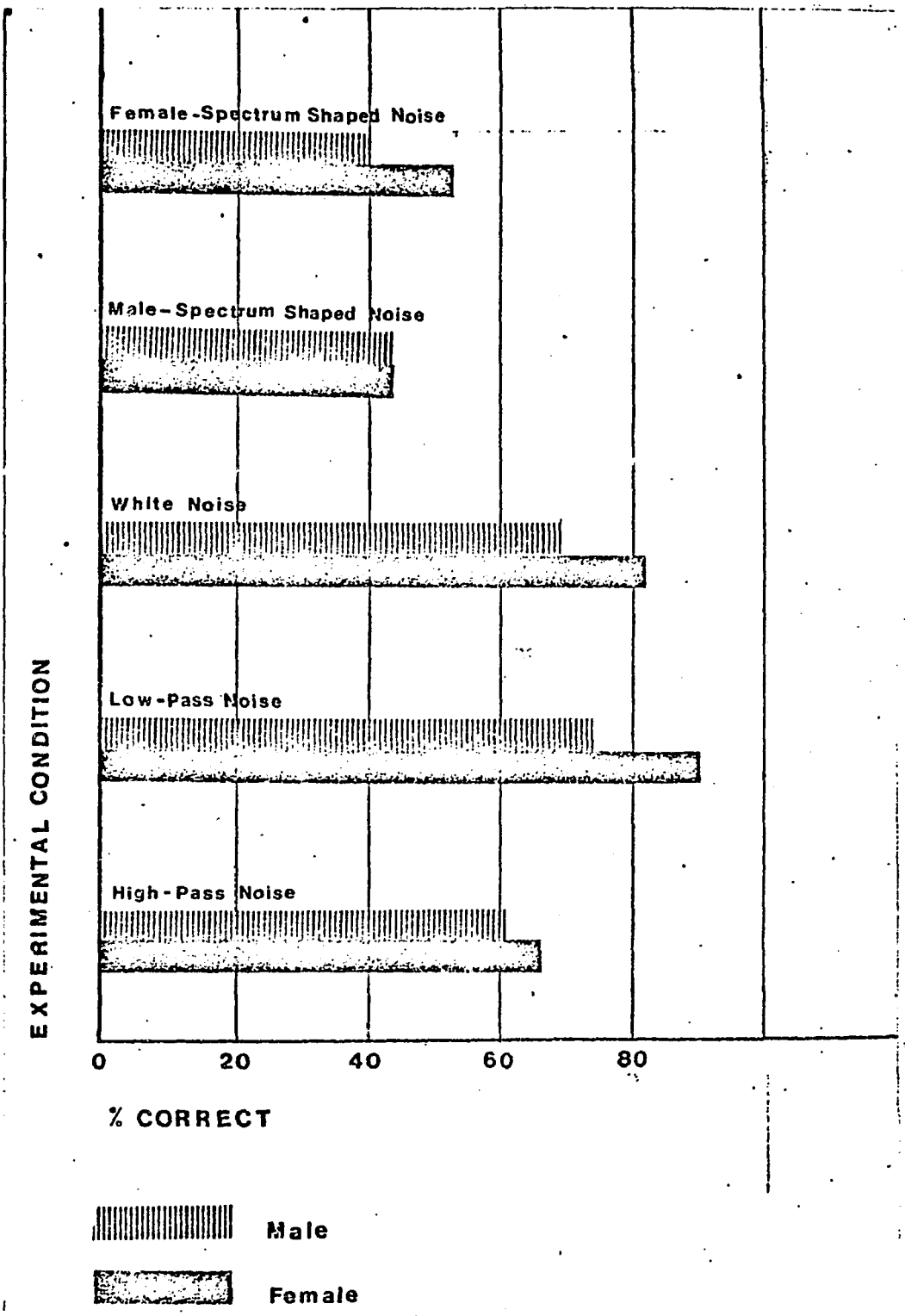
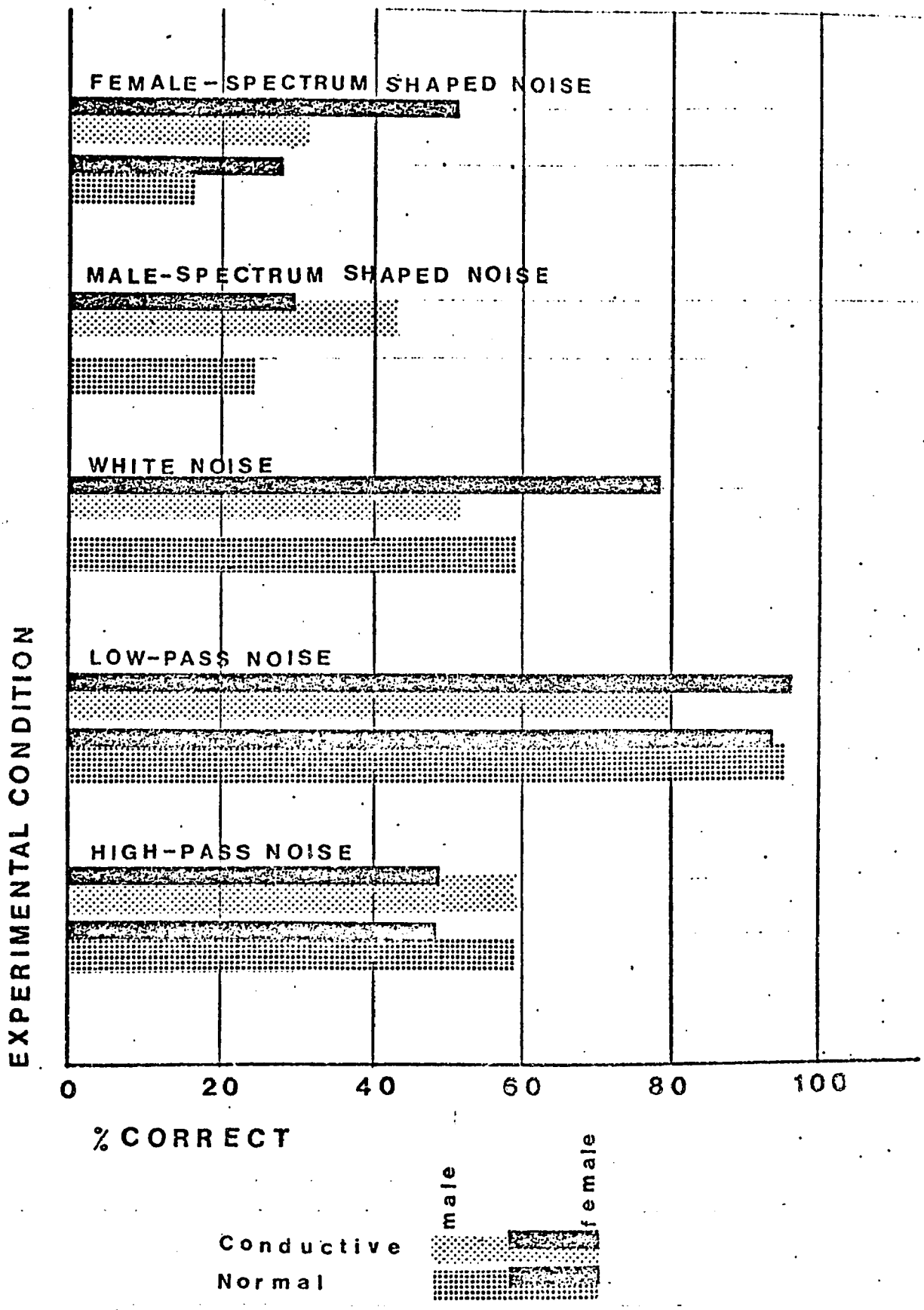

Female-Spectrum Shaped Noise
Male-Spectrum Shaped Noise
White Noise
Low-Pass Noise
High-Pass Noise
EXPERIMENTAL CONDITION
0
20
40
60
80
% CORRECT
Male
Female

Fig. 16. Mean percent scores for the experimental condition-experimental group interaction on the analysis of variance of the sentence data.


FEMALE-SPECTRUM SHAPED NOISE
MALE-SPECTRUM SHAPED NOISE
WHITE NOISE
LOW-PASS NOISE
HIGH-PASS NOISE
EXPERIMENTAL CONDITION
0
20
40
60
80
100
% CORRECT
male
female
Conductive
Normal

differed within each noise type. For example, for the white noise the listeners with the severely sloping sensori-neural hearing losses showed the highest scores whereas for the low-pass noise the listeners with the relatively flat hearing losses showed the highest scores.

Experimental Condition-Age Interaction. Figure 17 shows the average scores for each noise type for the two age groups. In all cases, the younger listeners obtained higher scores than the older listeners but the magnitude of this difference varied according to experimental condition. The smallest difference occurred with the white noise (1 percent).

Noise-Age-Experimental Group Interaction. Two three-level interactions were found to be statistically significant. Figure 18 shows the percent scores for each listener group and age group as a function of noise type for the first of these interactions. For the most part the younger listeners obtained better scores than the older listeners, a consistent exception being those listeners with conductive hearing losses and sometimes those with flat sensori-neural hearing losses. Also indicated is that, on average, the differences in score between the age groups increase for the male-spectrum and female-spectrum shaped noise conditions. The percent differences for the noise conditions are ranked as follows: white (5%), low-pass (11.5%),

Fig. 17. Mean percent scores for the experimental condition-age interaction on the analysis of variance of the sentence data.

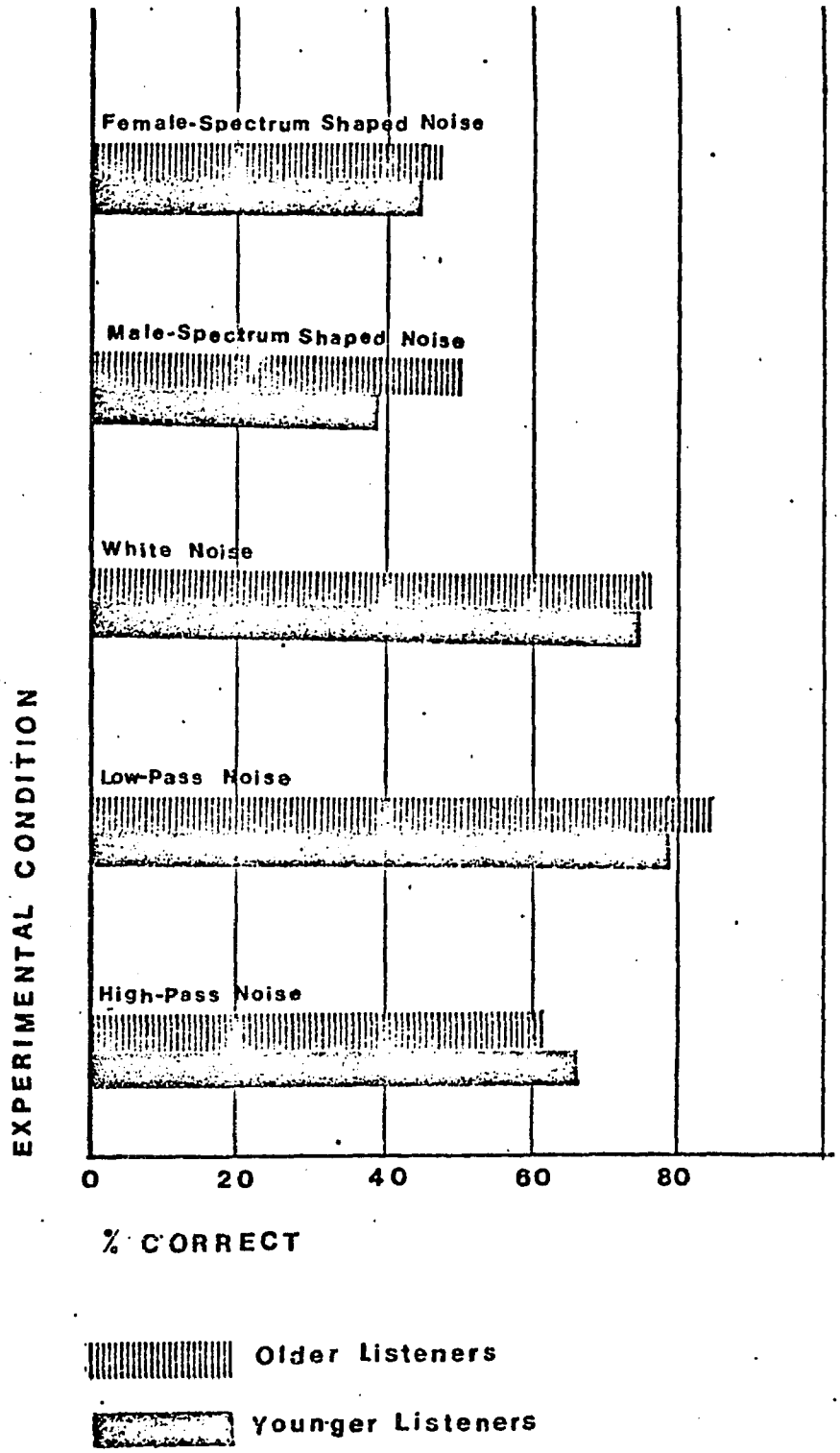
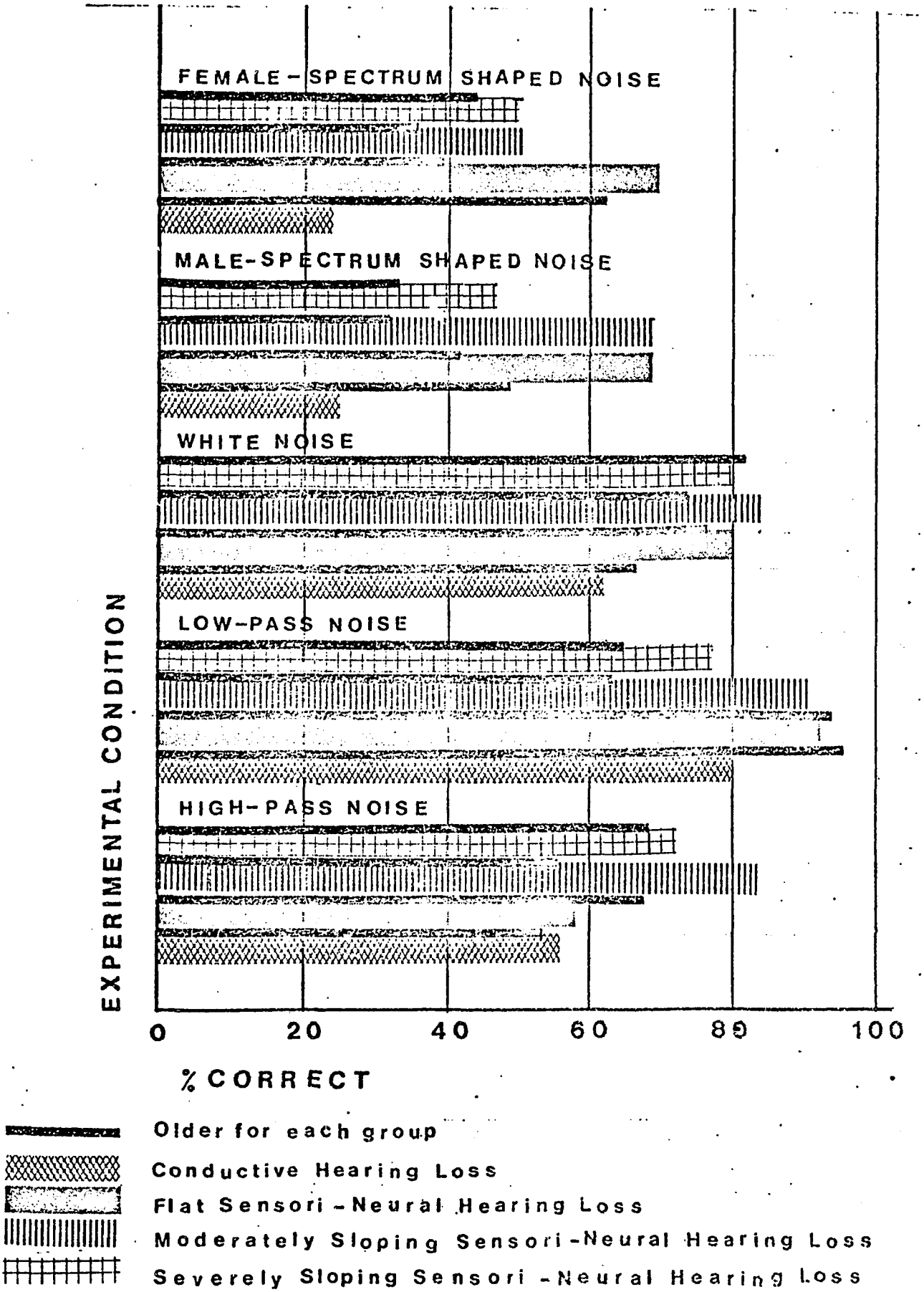

Female-Spectrum Shaped Noise
Male-Spectrum Shaped Noise
White Noise
Low-Pass Noise
High-Pass Noise
EXPERIMENTAL CONDITION
0
20
40
60
80
% CORRECT
Older Listeners
Younger Listeners

Fig. 18. Mean percent scores for the experimental condition-age-experimental group interaction on the analysis of variance of the sentence data.

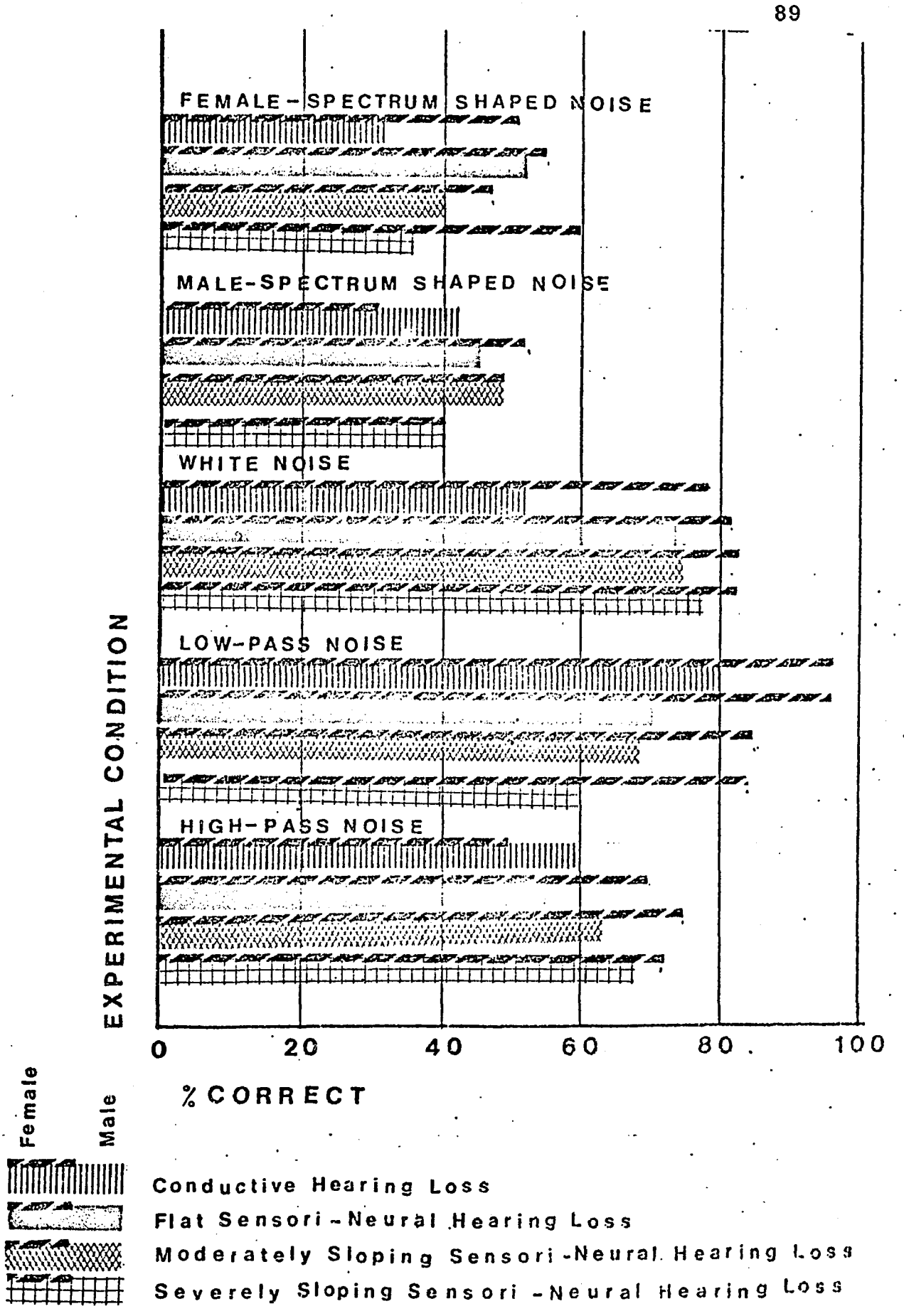

FEMALE-SPECTRUM SHAPED NOISE
MALE-SPECTRUM SHAPED NOISE
WHITE NOISE
LOW-PASS NOISE
HIGH-PASS NOISE
EXPERIMENTAL CONDITION
0
20
40
60
80
100
% CORRECT
Older for each group
Conductive Hearing Loss
Flat Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Severely Sloping Sensori-Neural Hearing Loss

high-pass (14.5%), female-spectrum shaped (22.8%) and male-spectrum shaped noise (25.8%).

Noise-Gender-Experimental Group Interaction. The second statistically significant three-level interaction was the Noise-Gender-Experimental Group interaction. These data are presented in Figure 19.

Those effects for Noise, Gender and Experimental Group described earlier are evident from this diagram. On average the intelligibility scores decreased systematically for the low-pass noise, white noise, high-pass noise, female-spectrum and male-spectrum shaped noise. Similarly there is an average decrease in intelligibility score in going from the flat to the severely sloping sensori-neural hearing loss group. The conductive listeners had a score comparable to that of the flat sensori-neural hearing loss group, but it should be remembered that the conductive hearing loss group had a S/N ratio of -15 dB for the sentence materials while the sensori-neural hearing loss groups had a 0 dB S/N ratio. Thirdly, in all but two cases the female talkers obtained higher intelligibility scores than the male talkers. However, the magnitude of this difference was greater for the low-pass noise, on average, as compared to the male-spectrum shaped noise (indicative of the interaction between Gender and Experimental Condition). The two conditions in which the male

Fig. 19. Mean percent scores for the experimental condition-gender-experimental group interaction on the analysis of variance of the sentence data.

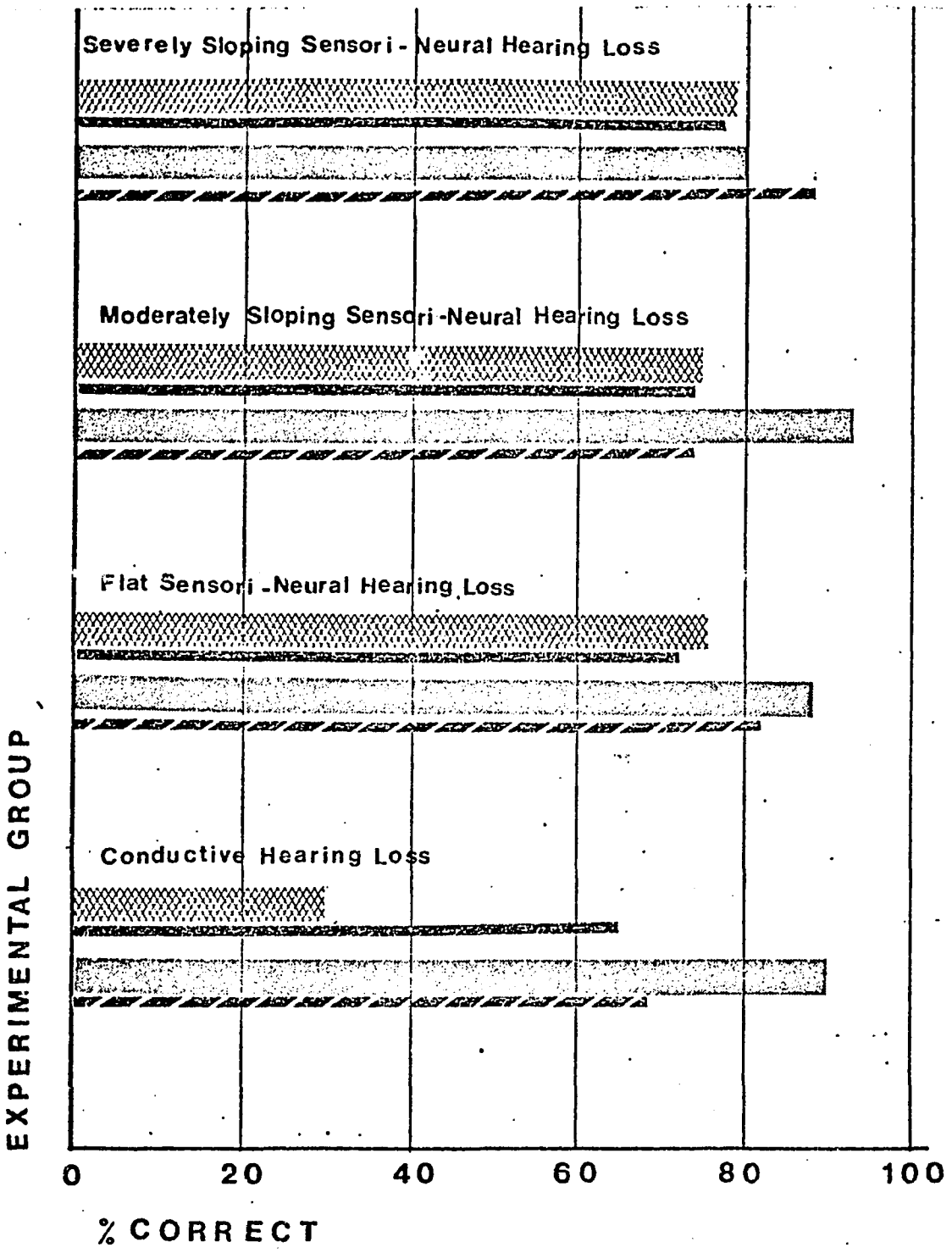

FEMALE-SPECTRUM SHAPED NOISE
MALE-SPECTRUM SHAPED NOISE
WHITE NOISE
LOW-PASS NOISE
HIGH-PASS NOISE
EXPERIMENTAL CONDITION
0
20
40
60
80
100
% CORRECT
Female
Male
Conductive Hearing Loss
Flat Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Severely Sloping Sensori-Neural Hearing Loss

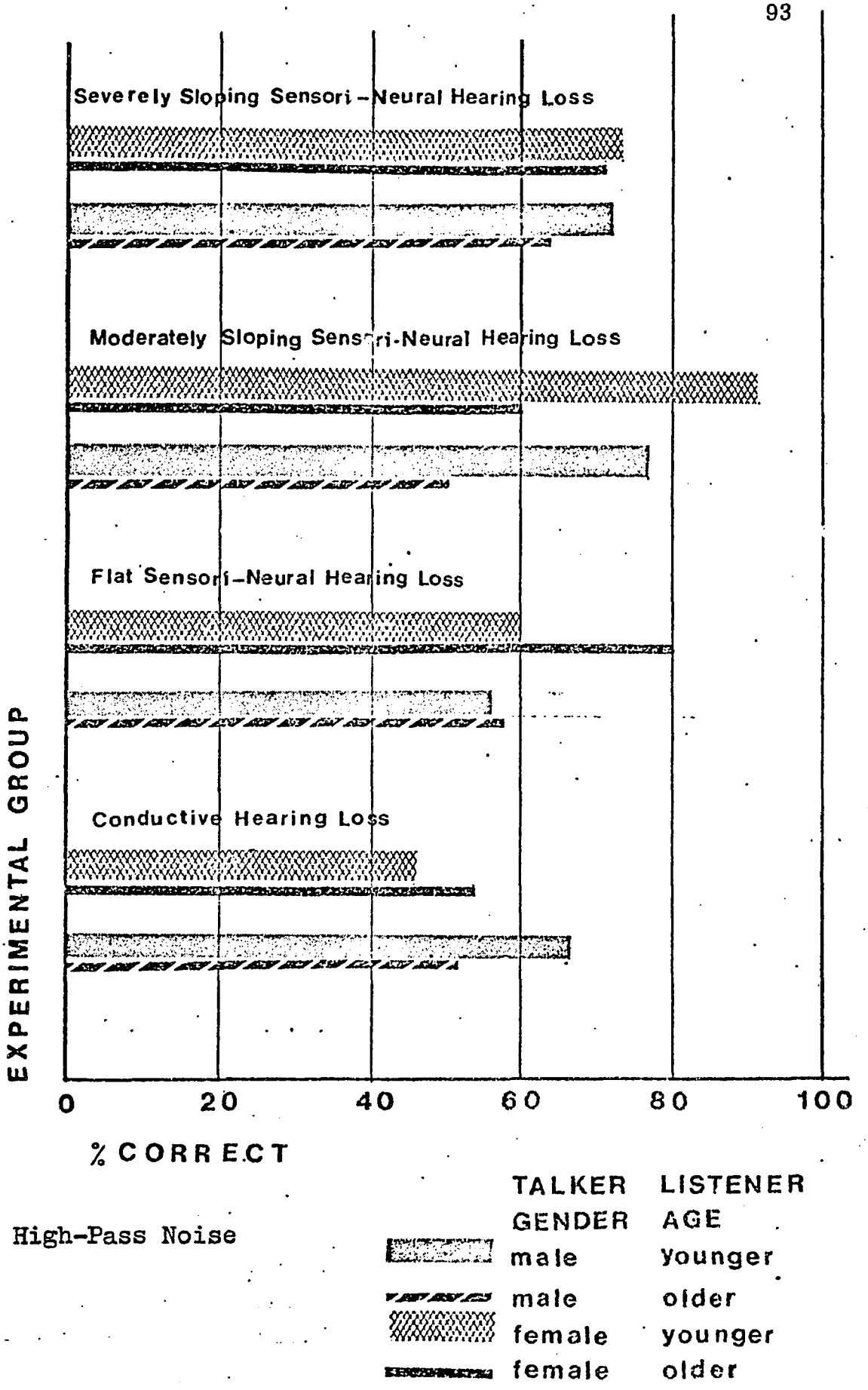
talkers obtained higher mean scores than the female talkers, occurred with the same Experimental Group (the conductive hearing loss group) but only for two Experimental Conditions (high-pass and male-spectrum shaped noise) is one aspect of the three-way interaction between these variables.

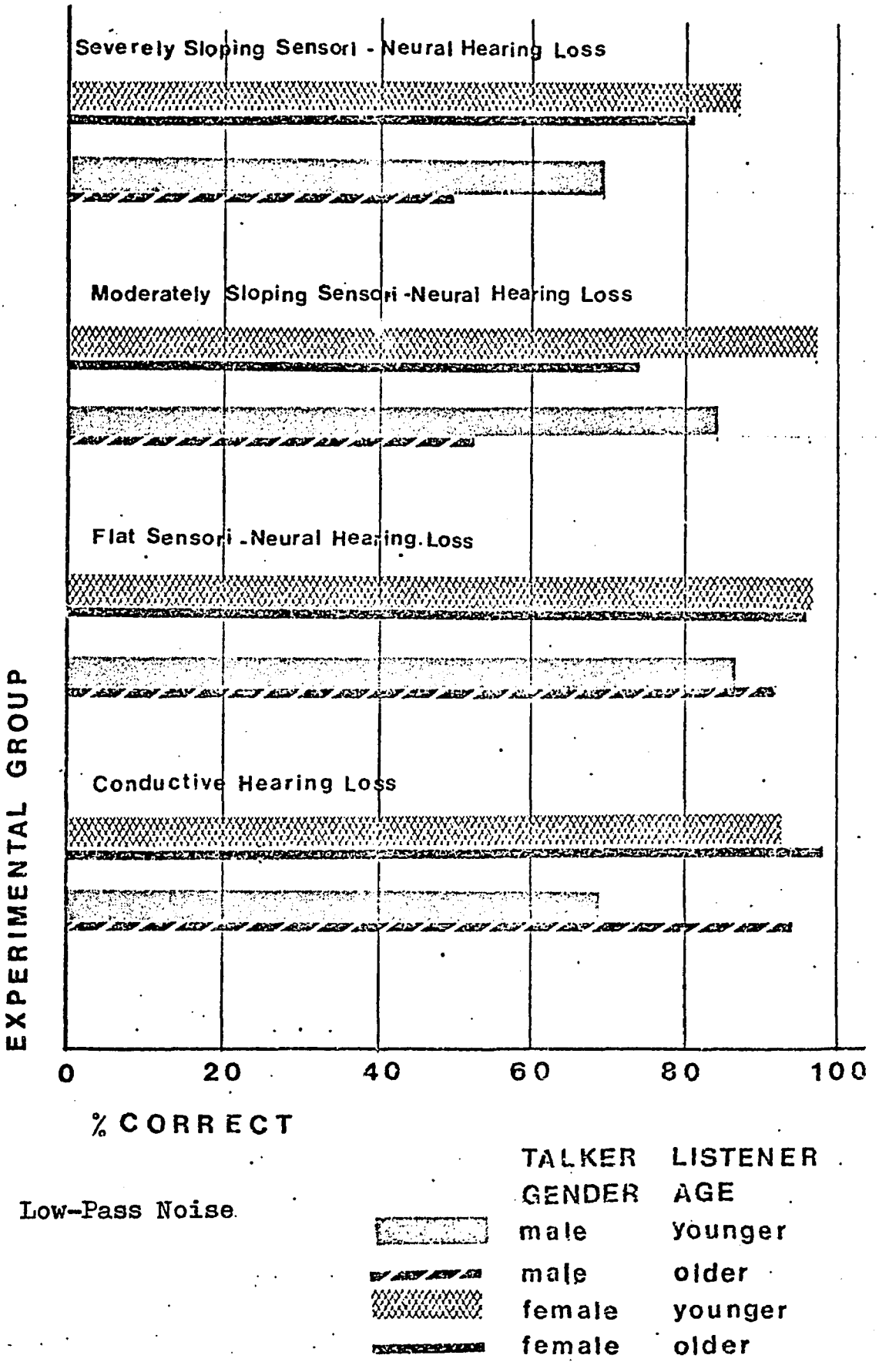
Noise-Gender-Age-Experimental Group Interaction. Only one four-level interaction was found to be statistically significant. Figures 20a-e show this interaction. As before, the main effects are evident in that the intelligibility scores decreased systematically for the different Experimental Conditions, the female talkers generally obtained higher intelligibility scores than the male talkers and there was a decrease in score within each Experimental Condition with variations per Experimental Group.

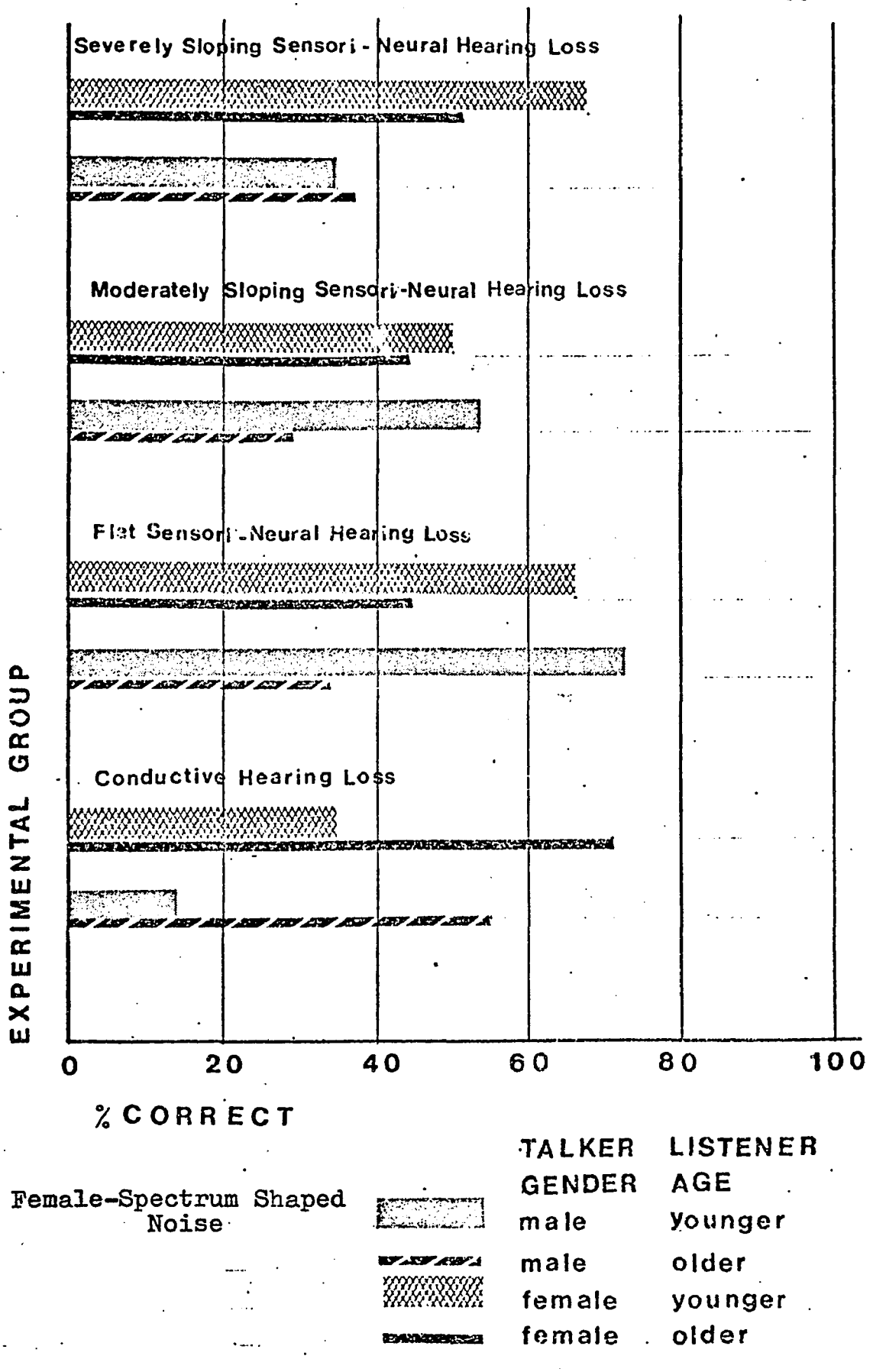
In addition, the figures show that there are a number of situations in which the older listeners obtained higher intelligibility scores than the younger listeners. Table 6 presents the age difference scores for both the male and female talkers for all Experimental Groups and Experimental Conditions. It is evident from this table that the reversal in expected results (older scoring better than younger listeners) occurred only for the conductive hearing loss group, but not for every Experimental Condition. In particular, higher scores were obtained for the younger conductively impaired

Figures 20a-e. Mean percent scores for the experimental condition-gender-age-experimental group interaction on the analysis of variance of the sentence data (a = white noise; b = high-pass noise; c = low-pass noise; d = female-spectrum shaped noise; e = male-spectrum shaped noise) of the hearing-impaired subjects.


Severely Sloping Sensori - Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Flat Sensori -Neural Hearing Loss
Conductive Hearing Loss
EXPERIMENTAL GROUP
0
20
40
60
80
100
% CORRECT
White Noise
TALKER GENDER
LISTENER AGE
male younger
male older
female younger
female older


Severely Sloping Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Flat Sensori-Neural Hearing Loss
Conductive Hearing Loss
EXPERIMENTAL GROUP
0
20
40
60
80
100
% CORRECT
High-Pass Noise
TALKER GENDER
LISTENER AGE
male younger
male older
female younger
female older


Severely Sloping Sensori - Neural Hearing Loss
Moderately Sloping Sensori -Neural Hearing Loss
Flat Sensori -Neural Hearing Loss
Conductive Hearing Loss
EXPERIMENTAL GROUP
0
20
40
60
80
100
% CORRECT
Low-Pass Noise
TALKER GENDER
LISTENER AGE
male younger
male older
female younger
female older


Severely Sloping Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Flat Sensori-Neural Hearing Loss
Conductive Hearing Loss
EXPERIMENTAL GROUP
0
20
40
60
80
100
% CORRECT
Female-Spectrum Shaped Noise
TALKER GENDER
LISTENER AGE
male younger
male older
female younger
female older

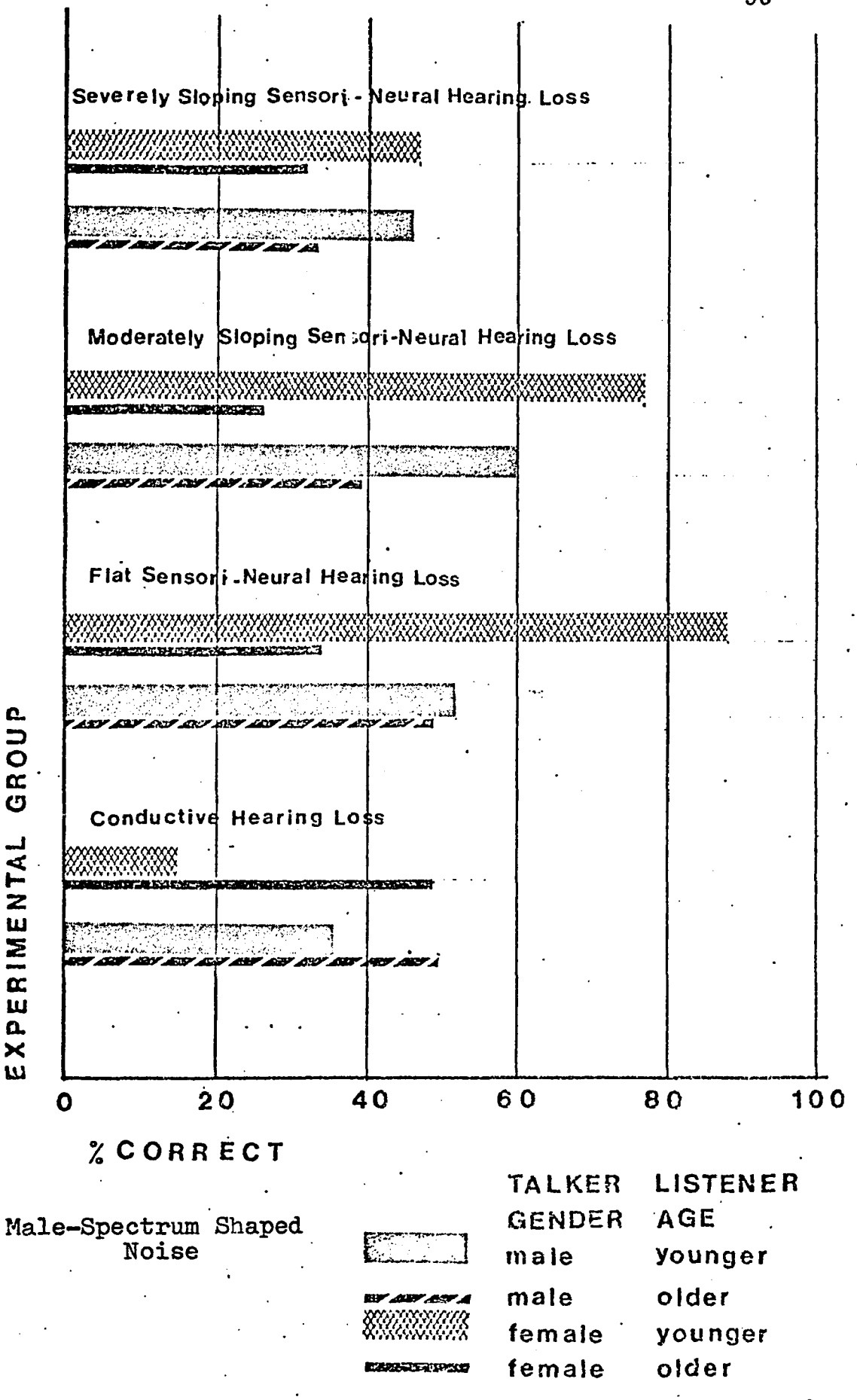

Severely Sloping Sensori - Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Flat Sensori-Neural Hearing Loss
Conductive Hearing Loss
EXPERIMENTAL GROUP
0
20
40
60
80
100
% CORRECT
Male-Spectrum Shaped Noise
TALKER GENDER
LISTENER AGE
male younger
male older
female younger
female older

TABLE 6

AGE DIFFERENCE SCORES FOR MALE AND FEMALE TALKERS FOR THE EXPERIMENTAL CONDITIONS AND EXPERIMENTAL GROUPS CONSIDERED*

	Experimental Group**										
	1		3		4		5		Mean		Mean
	Talker										
Noise	M	F	M	F	M	F	M	F	M	F	
High pass	14	-8	-1	-20	27	31	8	2	12	1.3	6.7
Low pass	-27	-5	-5	1	31	23	19	6	4.5	6.3	5.4
White	-27	22	3	2	1	19	2	-8	-5.3	8.8	1.8
Male-spectrum	-14	-34	3	54	21	51	13	15	5.8	21.5	13.7
Female-spectrum	-41	-36	39	22	24	6	-2	17	5	2.3	3.7
Mean	-19	-12.2	7.8	11.8	20.8	26	8	6.4	4.4	7.8	6.1
Mean	-15.6		9.8		23.4		7.2				6.2

*Difference scores were obtained by subtracting younger-older; negative number indicates the higher score was obtained by the older subjects.

**Group 1 indicates the conductive listeners.
Group 3 indicates flat sensori-neural listeners.
Group 4 indicates moderately sloping sensori-neural listeners.
Group 5 indicates severely sloping sensori-neural listeners.

listeners for the male talkers in the presence of high-pass noise and for the female talkers in the presence of white noise. In all other Experimental Conditions, the older conductive hearing loss group obtained higher intelligibility scores than the younger subjects.

Within the sensori-neural groups the least advantage for the two age groups was obtained with the severely sloping sensori-neural group where the scores were within seven percentage points. The largest age difference (23 percentage points) for these sensori-neural subjects was found for the moderately sloping group, the age difference for the listeners having a flat sensori-neural hearing loss was on the order of 10 percentage points which falls between the other two groups. With respect to Experimental Condition, the smallest age difference occurred for the white noise (1.8%), followed by the female-spectrum shaped noise (3.7%), low-pass noise (5.4%), high-pass noise (6.7%), and male-spectrum shaped noise (13.7%).

Sentence Intelligibility: normal hearing subjects

A separate analysis was performed on the data obtained from the normal hearing subjects because, as discussed earlier, the age range for these listeners was very small and no older subjects were tested.

In order to examine the data from the normal hearing listeners, an analysis of variance was performed between the sentence

scores of the listeners with conductive hearing losses and the normal hearing subjects, bearing in mind that the effects of age might not be the same for the two groups of listeners (mean age 25.9 years for the normal hearing listeners and 50.4 for the conductive hearing loss group).

The conductive hearing loss group was chosen because it was expected that there would be the greatest similarity between the two groups on the basis of discrimination ability. In addition, these two groups had the same S/N ratio (-15 dB) in contrast to the sensorineural groups which had a higher S/N ratio of 0 dB.

Table 7 presents the results of this analysis of variance between the normal hearing and conductive hearing loss subject groups. The results showed essentially the same type of findings as for the hearing-impaired listeners. Specifically, that Experimental Condition (masking noise), Talker Gender and the Noise-Gender interaction were statistically significant as well as the Experimental Group and its interactions with Noise and Noise-Gender.

The mean scores for the Noise, Talker Gender and Noise-Gender interactions for the two groups are presented in Table 8. An examination of these scores indicates that in most instances the scores for the two groups were very similar. The differences, while statistically significant, were not large. The conductive listeners obtained higher mean scores than the normal hearing listeners at an

TABLE 7

ANALYSIS OF VARIANCE FOR SENTENCE SCORES OF NORMAL AND CONDUCTIVE HEARING LOSS GROUPS OF LISTENERS

Source of Variation	Degrees of Freedom	Mean Squares	F-Ratio	Signif. Level
Noise (N)	4	.48	40.05	.001
Gender (G)	1	.01	1.25	.025
NG Interaction	4	.02	2.44	.005
Within subject variation	36	.01		
Experimental group (P)	1	.07	6.24	.005
NP Interaction	4	.02	2.40	.005
GP Interaction	1	.00	.73	
NGP Interaction	4	.01	1.42	.025
Between subject variation	144	.01		

TABLE 8

MEAN SENTENCE SCORES (PERCENT CORRECT) FOR NORMAL HEARING AND CONDUCTIVE HEARING LOSS GROUPS BY TALKER GENDER AND NOISE CONDITION

	Gender											
	Male						Female					
Noise	HP	LP	WN	MSN	FSN	Mean	HP	LP	WN	MSN	FSN	Mean
Normal hearing	58.7	95.4	58.5	24.3	15.6	50.5	48.6	93.9	51.9	27.0	27.6	49.8
Conductive hearing loss	58.9	80.3	51.7	42.9	31.5	53.1	49.5	95.5	78.8	29.7	51.7	61.0
Mean	58.8	87.9	55.1	33.6	23.6	51.8	49.1	94.7	65.4	28.4	39.7	55.4

HP - High pass noise
LP - Low pass noise
WN - White noise
MSN - Male-spectrum shaped noise
FSN - Female-spectrum shaped noise

equal S/N ratio (55.3% vs. 46.8%). Male talkers generally obtained lower scores than female talkers for both listener groups with the most noticeable exception occurring for the high-pass noise where the normal hearing group for the male talkers obtained a 10 percent higher mean score than the female talkers. Figure 21 presents the mean scores for the Noise-Gender-Experimental Group interaction.

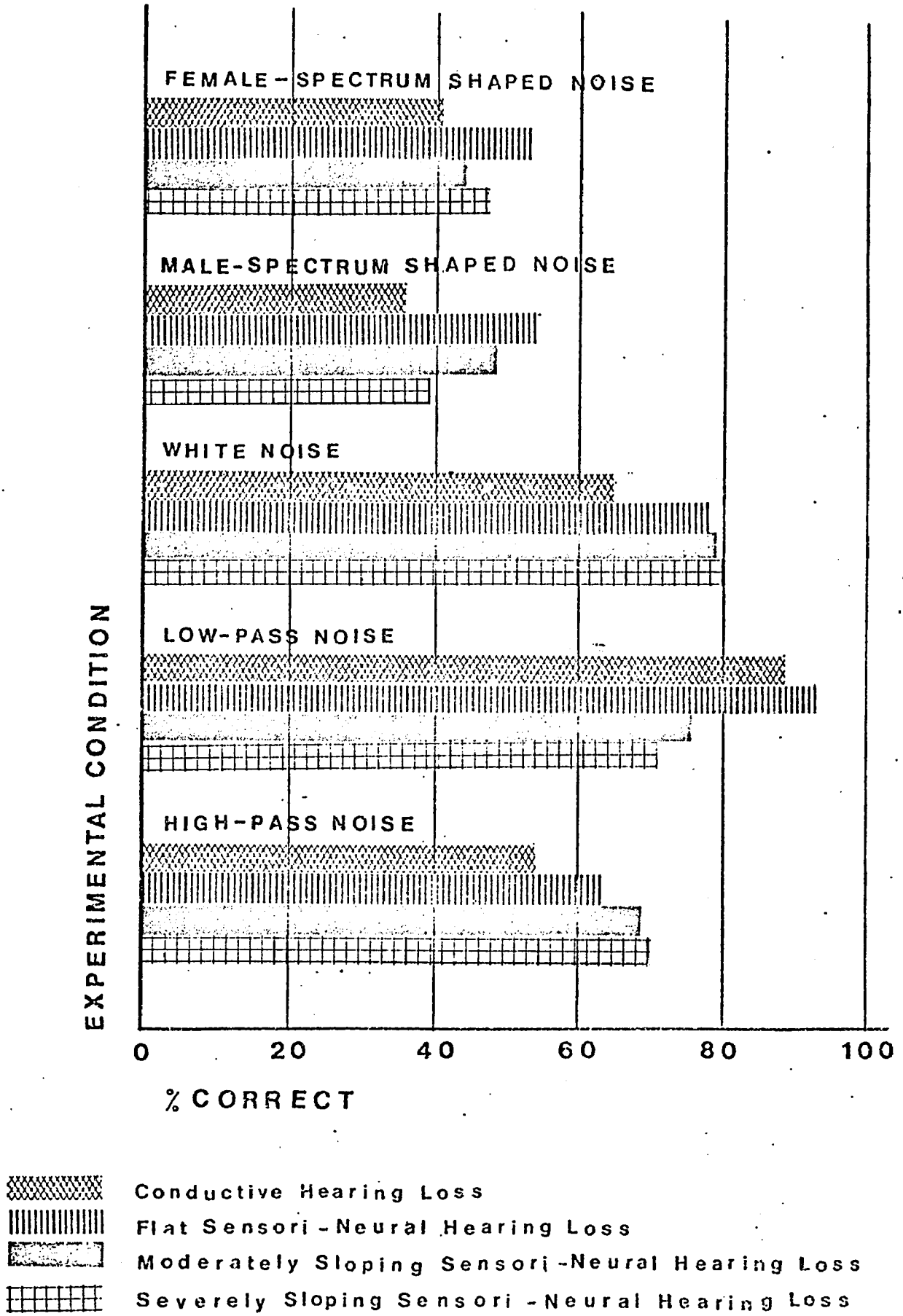
It is noteworthy that the masking effectiveness rankings presented earlier for the four hearing impaired listener groups remain constant when the normal hearing subjects are introduced: the low-pass noise being the least effective masker (94.6% mean score for the normal hearing subjects), followed by white noise (55.1%), high-pass noise (53.5%), male-spectrum shaped noise (25.6%), and female-spectrum shaped noise (21.2%).

Inter-talker differences

Since different talkers were used for different combinations of experimental conditions and groups, it was difficult to estimate inter-talker differences directly. However, it is possible for each combination of noise, talker and experimental group to be ranked on the basis of the relative performance of the different talkers.

For each combination of experimental group and condition three talkers were used per gender making it possible to rank the talkers in terms of relative intelligibility in sets of three. These

Fig. 21. Mean percent scores for the experimental condition-gender-experimental group interaction on the analysis of variance for the sentence data of the normal hearing and conductive hearing loss groups.


FEMALE-SPECTRUM SHAPED NOISE
MALE-SPECTRUM SHAPED NOISE
WHITE NOISE
LOW-PASS NOISE
HIGH-PASS NOISE
EXPERIMENTAL CONDITION
0
20
40
60
80
100
% CORRECT
Conductive Hearing Loss
Flat Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Severely Sloping Sensori-Neural Hearing Loss

rankings are shown in Table 9. The five Experimental Groups are represented in the columns and the five Experimental Conditions are represented in the rows. The six entries per cell show the identities of the three male and three female talkers ranked in order of intelligibility.

Certain trends are evident from the table. Of the male talkers, M5 showed the highest intelligibility scores in almost every case. M4 was consistently the next most intelligible talker (and best in those cells where M5 was not represented), M1, M2, M3 showed similar intelligibility scores, and their relative rankings varied between conditions. Of the female talkers, F2 consistently obtained the highest intelligibility scores. F3 and F4 showed similar rankings, consistently below F2 and above both F1 and F5. F1 and F5 showed essentially the same rankings.

It is interesting to note that while the females obtained higher intelligibility scores, on average, the best male talker was consistently better than at least two of the female talkers (F1 and F5) and roughly comparable to F3 and F4. The lowest intelligibility scores for both the male and female talkers were about equal. In comparing male and female talkers the female talkers were ranked higher, on average, but the two groups overlapped consistently. In only one of the fifteen conditions considered, the low-pass noise condition for the severely sloping sensori-neural listeners, the female talkers are

TABLE 9

TALKER RANKINGS IN ORDER OF INTELLIGIBILITY SCORES (WITH SIX ENTRIES PER CELL REPRESENTING THE SIX TALKERS PER CONDITION)

Noise	Talker	Experimental Group* 1	2	3	4	5
High-pass	M	M4M3 M1	M4M3 M1	M1M3M4	M4 M3 M1	M3 M4 M1
	F	F1F5F3	F5F1F3	F3F1F5	F3F5 F1	F3 F1F5
Low-pass	M	M3 M2 M1	M1M2M3	M1 M3M2	M1M2 M3	M2 M3M1
	F	F3F5 F4	F3F4 F5	F3F4 F5	F4F3 F5	F3F5F4
White	M	M5M1 M2	M5 M2M1	M5M2 M1	M5M2 M1	M5 M2 M1
	F	F2F3 F1	F2 F3 F1	F2 F3F1	F2 F1 F3	F2 F3 F1
Male-spectrum	M	M4M2M5	M5 M4M2	M2M5M4	M5M4 M2	M5M4 M2
	F	F2 F4F1	F2 F4 F1	F2F4 F1	F2 F1F4	F2 F4 F1
Female-spectrum	M	M5 M4M3	M5M4 M3	M5M4 M3	M5 M3M4	M5 M4M3
	F	F4 F2F5	F2 F4F5	F2 F4F5	F5F4F2	F4F2 F5

*Numbers refer to group assignments used throughout the study (see Table 6).

more intelligibile than the male talkers. In one other case, high-pass noise for the flat sensori-neural group, there was marginal overlap in that the lowest ranked female talker was equal in intelligibility to the highest male talker. In all other conditions one or more male talkers were ranked above one or more female talkers in terms of relative intelligibility.

An additional factor affecting the relative rankings between the male and female talkers is the spectrum of the masking noise. When the high-pass noise is presented the male talkers score higher (more intelligible) than the female talkers. When the masking signal is white noise the female talkers are more intelligible. When low-pass noise is presented the effect is the same. For male-spectrum shaped noise the female talkers are better, but for the female-spectrum shaped noise the results are mixed.

In summary, there appears to be a consistent pattern showing large differences between talkers, but the female talkers, on the average, were more intelligible than the male talkers in the study. In particular, the best female (F2) is consistently more intelligible than the best male (M5). In addition, it was determined that the spectrum of the masking noise affects the relative differences between the male and female talkers.

Summary: Sentence Intelligibility for all subjects

The following effects were found to be significant for all listeners whether or not they were hearing impaired:

1. The female talkers were, on average, more intelligible than the male talkers (even after the adjustment for the low signal-to-noise ratio for one male talker).

2. There are consistent masking noise effects with the different maskers producing different amounts of intelligibility decrement. The low-pass noise was the least effective masker, followed by white noise, high-pass noise, and finally male-spectrum and female-spectrum shaped noise.

3. For the hearing-impaired subjects, intelligibility scores for the groups decreased as the high frequency sensori-neural hearing loss increased. Age was also found to be a significant factor, but an unexpected interaction was noted for the conductive hearing loss group where the older subjects consistently scored higher than their younger counterparts.

Syllable Discrimination: subjects with hearing loss

The same sequence of analyses was followed for the nonsense syllables as was used for the sentences. These analyses were then followed by an analysis of the error patterns obtained with the non-

sense syllables. In all analyses involving the nonsense syllables, the percent correct score is used. Any adjustments for random guessing effects were not included.

The first analysis of variance considered the following factors: Experimental Condition, Talker, Age and Experimental Group. The method used for testing statistical significance for between and within subjects was the same as was used in the earlier sections. Also as discussed earlier, the normal hearing listeners were not included in the main analysis of variance in order that age might be one of the variables.

As can be seen in Table 10, all of the main effects were found to be statistically significant: Experimental Condition, Age, Gender and Experimental Group. These factors are discussed below.

Experimental Condition. The four experimental conditions showed large differences in mean score. The highest scores were obtained for the +6 dB/octave slope condition. Scores for the 0 dB/octave slope and the -6 dB/octave slope conditions were very similar and there was a very large drop in the scores for the noise condition. The mean scores for each of the Experimental Conditions are shown in Table 11.

Talker. Large talker differences were obtained with the male talker obtaining a 66.1 percent mean score as compared to a

TABLE 10

ANALYSIS OF VARIANCE FOR NONSENSE SYLLABLE SCORES OF HEARING-IMPAIRED LISTENERS

Source of Variation	Degrees of Freedom	Mean Squares	F-Ratio	Signif. Level
Experimental condition (F)	3	7.52	483.20	.001
Talker (G)	1	.55	35.54	.001
FG Interaction	3	.07	4.82	.007
Within subject variation	99	.01		
Age (A)	1	.65	5.70	.005
Experimental group (P)	3	3.27	28.68	.001
AP Interaction	3	1.00	8.77	.001
FA Interaction	3	.00	.05	
GA Interaction	1	.08	7.39	.001
FGA Interaction	3	.02	.21	
FP Interaction	9	.09	.82	
GP Interaction	3	.00	.04	
FGP Interaction	9	.01	.16	
FAP Interaction	9	.14	1.28	.001
GAP Interaction	3	.02	.19	
FGAP Interaction	9	.01	.09	
Between subject variation	220	.11		

TABLE 11

NONSENSE SYLLABLE SCORES FOR HEARING IMPAIRED LISTENERS FOR FOUR EXPERIMENTAL CONDITIONS

Experimental Condition	Percent Correct
0 dB/octave slope	68.3
+6 dB/octave slope	77.4
-6 dB/octave slope	66.8
0 dB/octave slope + noise	42.8

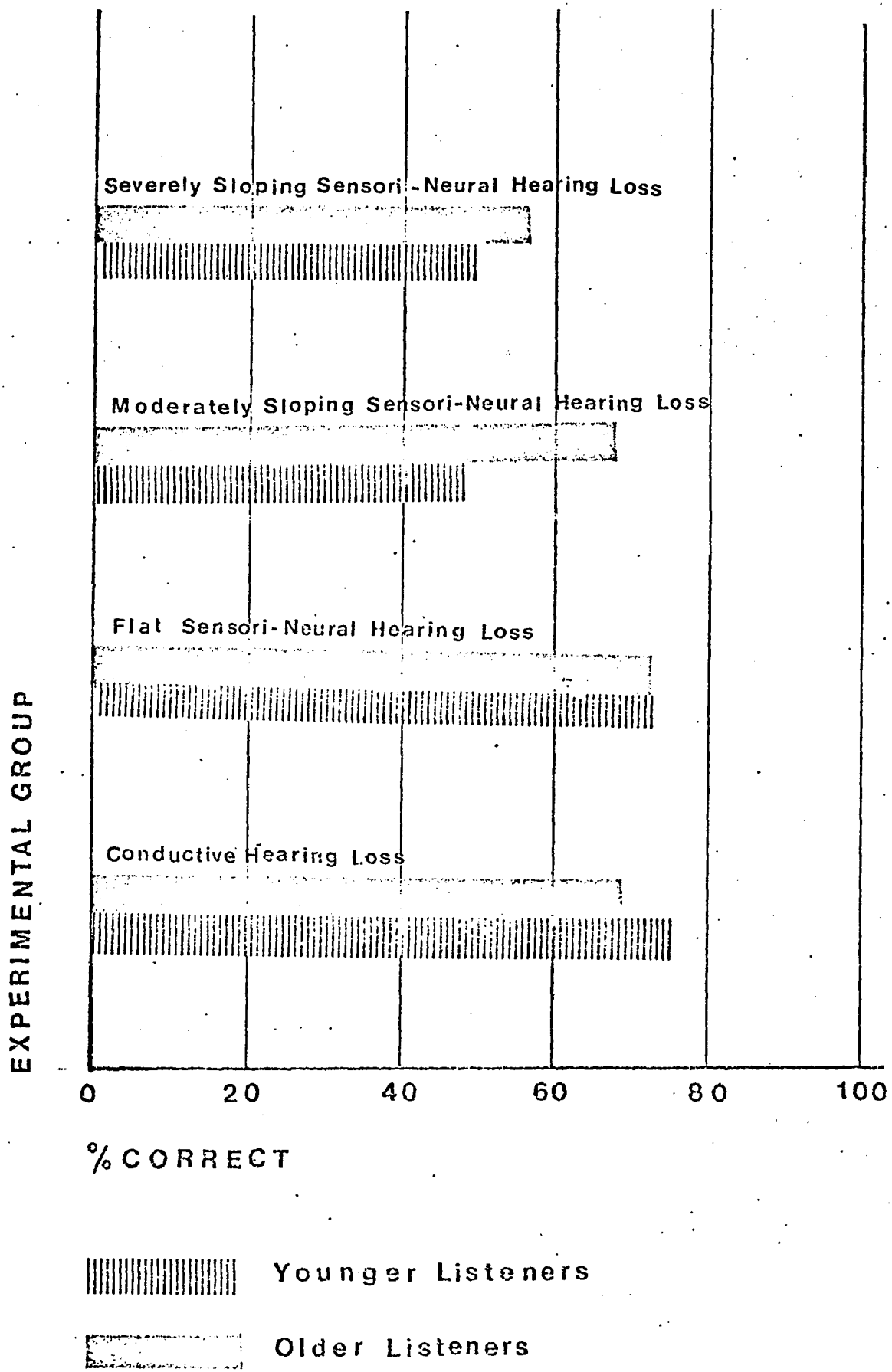
62.2 percent mean score for the female talker when averaged across all conditions. In comparison, on the sentence data, this same male talker obtained a higher score than the corresponding female talker for the same noise conditions.

Age. The effect of age on the nonsense syllable materials was of approximately the same magnitude as the effect of talker gender. The younger listeners obtained higher scores than the older subjects, averaged over both Experimental Group and Experimental Condition. The mean score for the younger listeners was 66.4 percent as compared to a 62.0 percent for the older subjects.

Experimental Group. The average scores for the four experimental groups considered in this analysis show a similarity based upon the audiological configuration of the hearing loss. The mean scores for the conductive and flat sensori-neural groups were almost identical (72.0% vs. 72.5%). The moderately sloping and severely sloping sensori-neural groups obtained scores markedly below the previous groups but with similar mean scores of 57.4% and 54.0%, respectively.

Age-Experimental Group Interaction. The results of the Age-Experimental Group Interaction are presented in Figure 22. As can be seen in the figure, there are differences in the performance

Fig. 22. Mean percent scores for the hearing impaired listeners for the age-experimental group interaction on the analysis of variance of the nonsense syllable data.


Severely Sloping Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Flat Sensori-Neural Hearing Loss
Conductive Hearing Loss
EXPERIMENTAL GROUP
0
20
40
60
80
100
% CORRECT
Younger Listeners
Older Listeners

of the conductive and flat sensori-neural groups compared to the two sloping sensori-neural groups. The older listeners with the flat losses obtained higher scores than their younger counterparts, although the differences are not large. This situation is reversed for the two sloping sensori-neural groups where the younger subjects obtained the higher scores and the magnitude of the mean differences increased markedly from 2% to 5% for the flat configurations to 7% to 18% for the sloping groups.

Gender-Age Interaction. Figure 23 presents the results of the Gender-Age interaction. In this comparison the male talker obtained the higher mean score for both the younger and older listeners, although the difference between the groups for the female talker is of smaller magnitude than the male talker.

Experimental Condition-Gender Interaction. Figure 24 shows the scores for the male and female talker on each experimental condition. On average, the scores for each condition show a systematic decline going from the +6 dB/octave slope filter condition to the 0 dB/octave slope filter condition, to the -6 dB/octave slope condition, to the noise condition. The scores for the male talker are consistently higher than those for the female, but the magnitude of the difference changes between Experimental Conditions. For example, it is only 2.5 percentage points for the -6 dB/octave slope condition, whereas

Fig. 23. Mean percent scores for the gender-age interaction on the analysis of variance of the nonsense syllable data of the hearing impaired listeners.

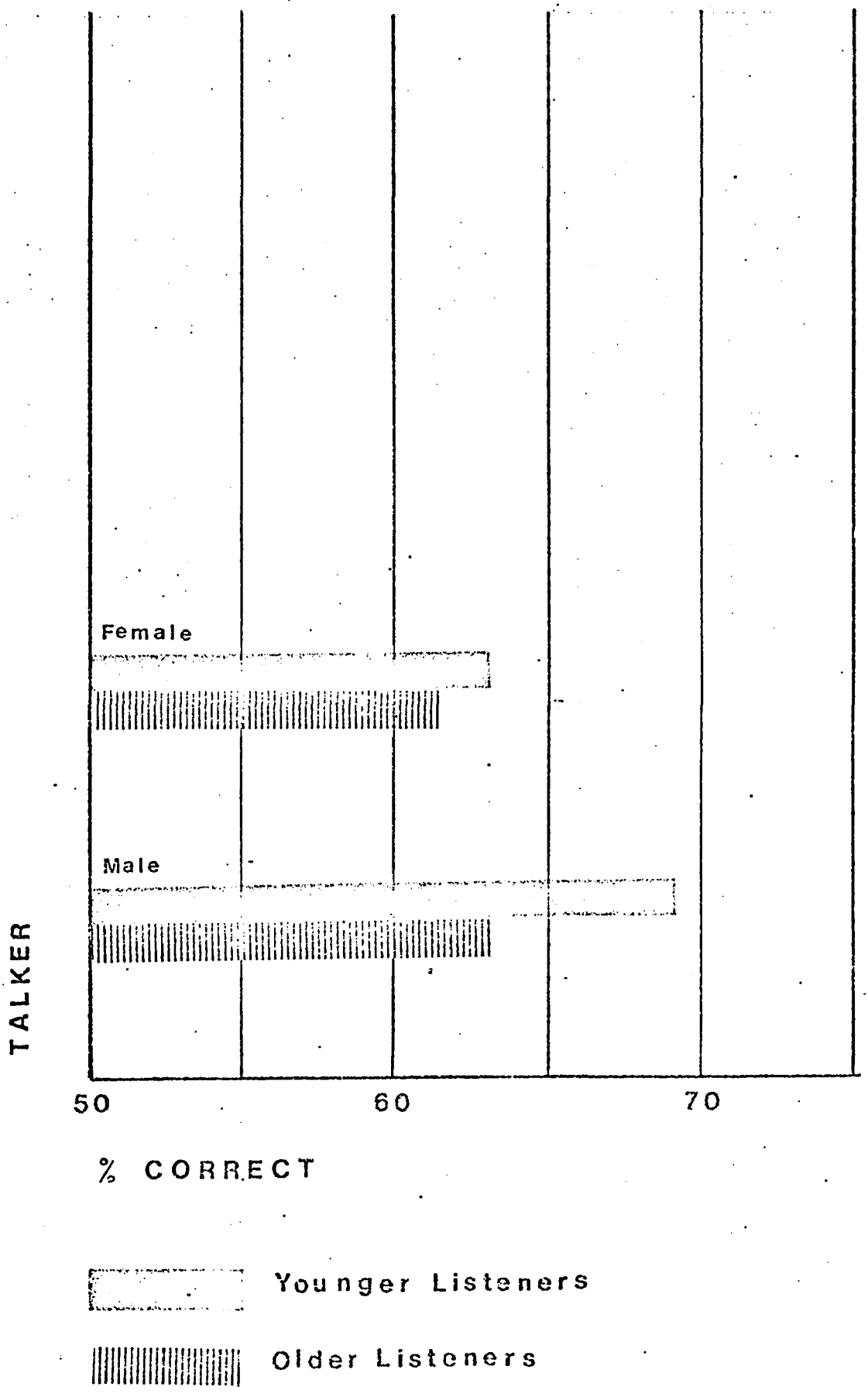
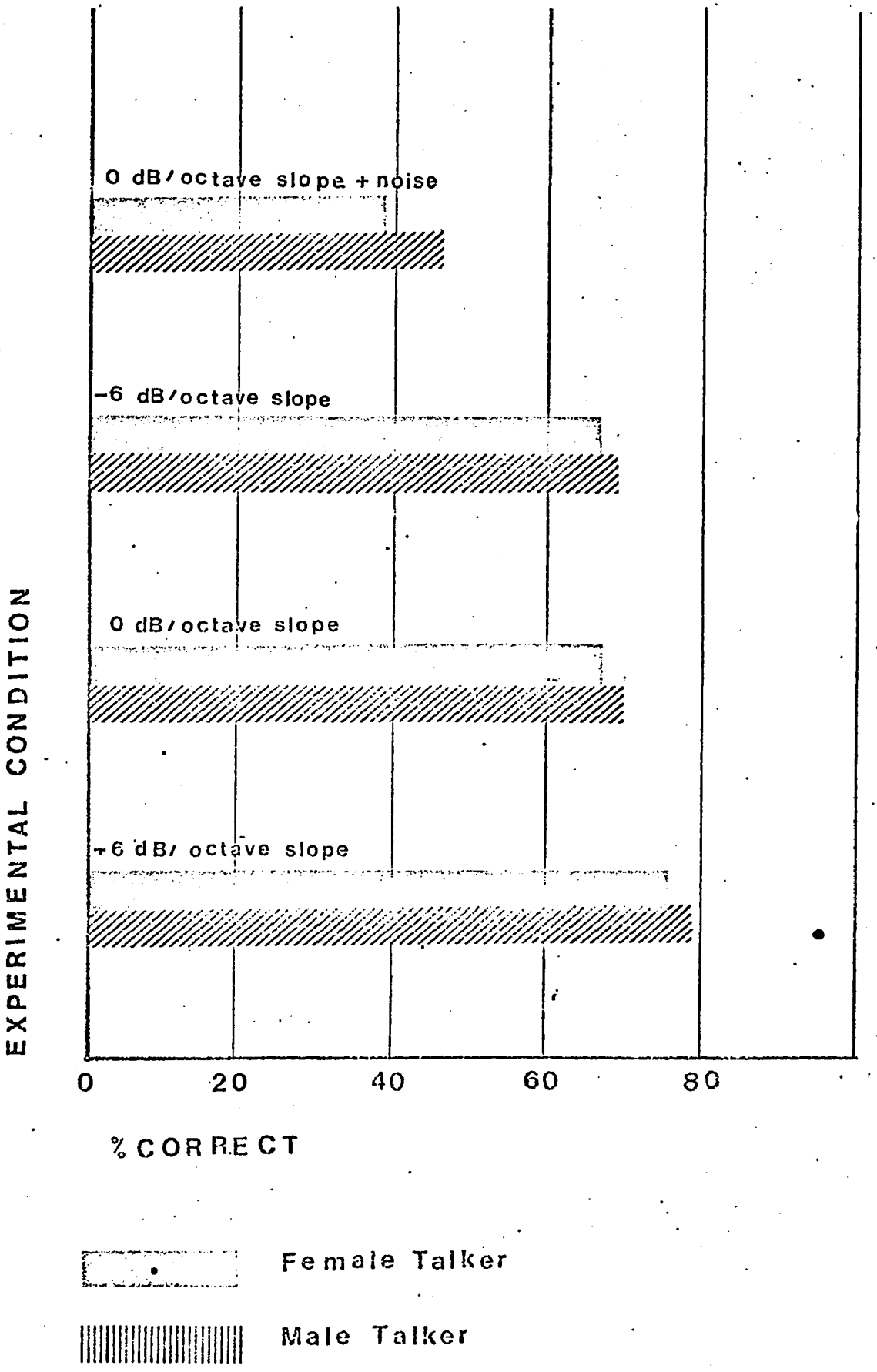

Female
Male
TALKER
50
60
70
% CORRECT
Younger Listeners
Older Listeners

Fig. 24. Mean percent scores for the experimental condition-gender interaction for the analysis of variance on the nonsense syllable data of the hearing impaired listeners.


0 dB/octave slope + noise
-6 dB/octave slope
0 dB/octave slope
+6 dB/ octave slope
EXPERIMENTAL CONDITION
0
20
40
60
80
% CORRECT
Female Talker
Male Talker

for the noise condition, the score for the male talker is over 9 percent higher than that of the female talker.

Experimental Condition-Age-Experimental Group Interaction. Only one three-level interaction was found to be statistically significant. Figure 25 shows the percent scores for each listener and age group as a function of experimental condition. For the most part the younger listeners obtained better scores than the older listeners, with the exception being the conductive hearing loss group. Also indicated is that the range of scores within experimental conditions was large for the -6 dB/octave slope condition, and relatively small for the +6 dB/octave slope and spectrum-shaped noise conditions.

Syllable Discrimination: normal hearing subjects

A separate analysis of variance was used in analyzing the data obtained for the normal hearing listeners. As was done in the analysis of the sentence materials, both normal hearing and conductively impaired subjects were included in this analysis because there was reason to expect similarities in the performance of the two groups.

Table 12 shows the results of the analysis of variance. The results of the analysis show that both of the factors which are varied within each subject are statistically significant, namely Experimental

Fig. 25. Mean percent scores for the experimental condition-age-experimental group interaction for the analysis of variance of the nonsense syllable data of the hearing impaired listeners.

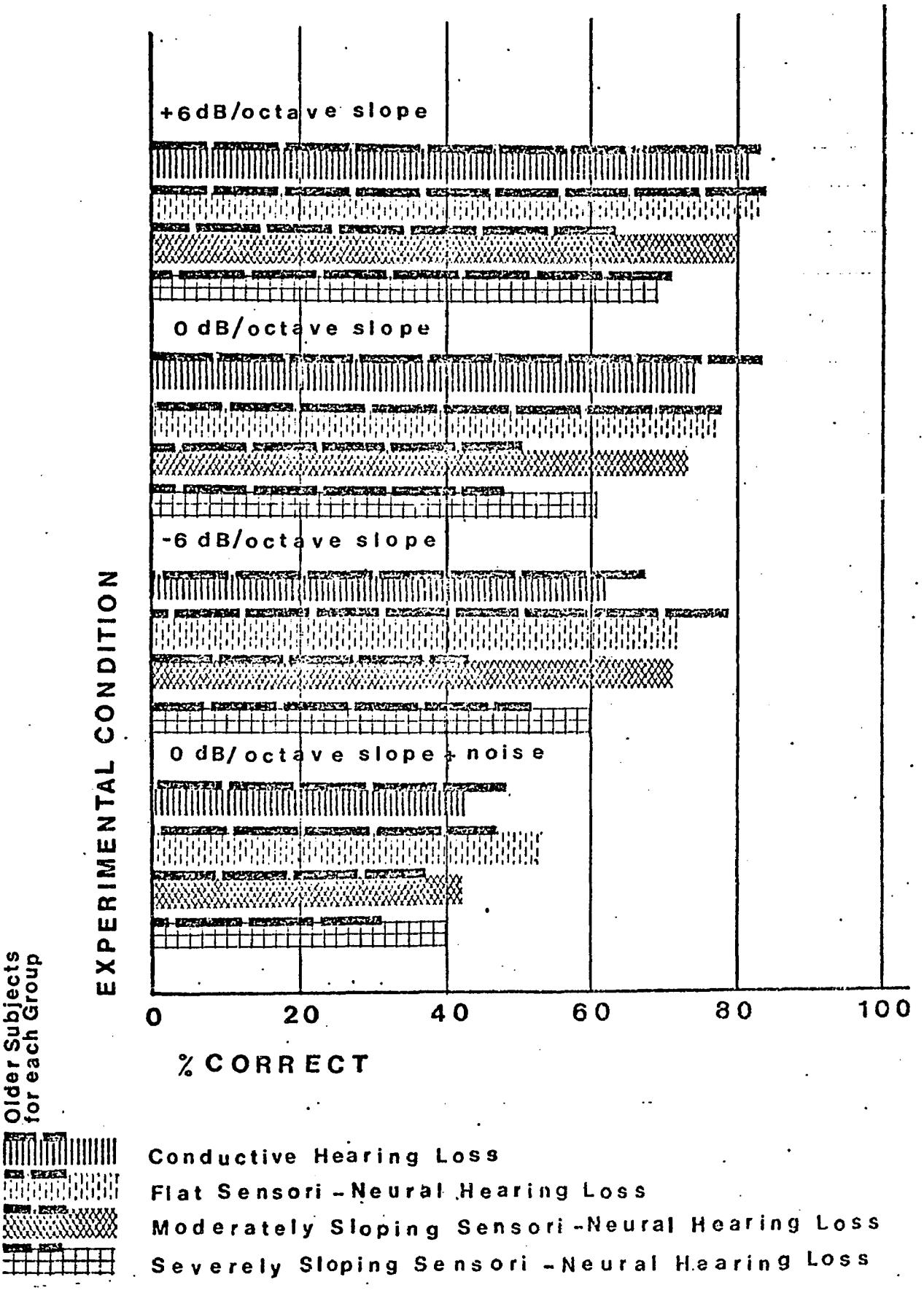

+6dB/octave slope
0 dB/octave slope
-6 dB/octave slope
0 dB/octave slope + noise
EXPERIMENTAL CONDITION
0
20
40
60
80
100
% CORRECT
Older Subjects for each Group
Conductive Hearing Loss
Flat Sensori-Neural Hearing Loss
Moderately Sloping Sensori-Neural Hearing Loss
Severely Sloping Sensori-Neural Hearing Loss

TABLE 12

ANALYSIS OF VARIANCE ON NONSENSE SYLLABLE MATERIALS FOR NORMAL HEARING AND CONDUCTIVE HEARING LOSS GROUPS OF LISTENERS

Source of Variation	Degrees of Freedom	Mean Squares	F-Ratio	Signif. Level
Experimental condition (F)	3	10989.0	620.	.001
Talker (G)	1	743.9		.001
FG Interaction	3	106.8	6.03	.003
Within subject variation	17	17.7		
Experimental group (P)	1	6088.5	3.48	.010
FP Interaction	3	94.9	.05	
GP Interaction	1	.1	.00	
FGP Interaction	3	7.5	.00	
Between subject variation	117	1747.5		

Condition and Talker effects. Of the factors varied across subjects, Experimental Group was found to be statistically significant.

Since the effect of hearing loss was shown to be significant, the mean scores for the two groups were then examined. Those scores are presented in Table 13. The normal hearing listeners obtained higher mean scores (82.4%) than the conductive hearing loss group, the latter obtaining a mean score of 70.0%. Upon comparison with the mean scores of the sensori-neural groups, it was found that the conductive and flat sensori-neural groups yielded essentially the same score (71%), and the scores for the sloping sensori-neural groups decreased from 57% to 53% as the high frequency involvement increased. As was the case for the sentence materials, the normal-hearing subjects obtained the highest scores of any of the Experimental Groups.

Summary: nonsense syllable discrimination

The analyses on the nonsense syllable scores indicated that:

1. The +6 dB/octave slope filter condition provides the greatest intelligibility for all listeners of all ages.

2. Masking by speech-spectrum shaped noise produced the lowest syllable discrimination scores for all listeners of all ages.

3. The male talker was consistently more intelligible than the female talker for all Experimental Groups, Experimental

TABLE 13

MEAN NONSENSE SYLLABLE PERCENT SCORES FOR NORMAL AND CONDUCTIVE HEARING LOSS GROUPS BY TALKER AND EXPERIMENTAL CONDITION

	Talker									
	Male					Female				
Experimental Condition	+6	-6	0	Noise	Mean	+6	-6	0	Noise	Mean
Normal hearing	95.7	92.0	89.9	60.4	84.5	93.5	88.8	87.1	51.6	80.3
Conductive hearing loss	82.5	76.9	77.7	51.8	72.2	79.3	72.2	77.3	42.6	67.9
Mean	89.1	84.5	83.8	56.1	93.8	86.4	80.5	82.2	47.1	74.1

Conditions and Ages.

4. The flat audiometric configuration obtained higher scores than either of the sloping sensori-neural groups.

Effects on individual nonsense syllables

In order to examine the effects of the experimental variables on individual syllables of the nonsense syllable test, confusion matrices were obtained for each combination of the main factors, i.e., Experimental Group, Talker, Experimental Condition and Age. The rows of each confusion matrix represent the intended syllables (i.e., targets) and the columns identify the syllable intended by the listener (i.e., responses).

Two analyses were performed. The first dealt with the frequency of correct identification of each syllable. These data were provided by the proportions in the diagonal cells of the matrix. The second analysis dealt with the confusions, i.e., entries in cells other than the diagonal. The analysis of correct identification will be described first.

Analyses of variance were performed on the target responses of each matrix. Sets of four factors were considered for each subtest: Syllable, Talker, Experimental Condition and Experimental

Group (or Age).[3]

A separate analysis of variance was done for each subtest since the target syllables differed between subtests. Table 14 presents the results of these for each subtest of the nonsense syllable test. As can be seen in Table 14, there are certain factors that consistently show statistically significant differences. These were Target Syllable, Experimental Condition and Experimental Group. Of the interactions considered, only Target Syllable-TalkerGender, Target Syllable-Experimental Condition and Target Syllable-Experimental Group interactions consistently show statistically significant effects.

Of primary interest in this section are the Target Syllable effects and any interactions involving the target syllables; the effects of the other variables in the analysis have been discussed earlier. Table 15 presents the mean scores for each target in the analysis. In this table the rows represent the target syllable and the columns represent the percent correct for each target for each individual subtest. The table separates the targets into voiced and voiceless consonants and reveals several trends regarding the target syllables:

1. The syllables with voiced consonants typically showed

---

[3]It was not possible to compare Experimental Group and Age variables for all combinations of these two variables since the normal-hearing group consisted of young listeners only.

TABLE 14

SIGNIFICANCE LEVELS FOR THOSE EFFECTS INVOLVING DIFFERENT TARGETS OF THE NONSENSE SYLLABLE SUBTESTS CONSIDERED BY EXPERIMENTAL GROUP

Source of Variation	Subtest 1	2	3	4	5	6	7
Target (T)	.001	.001	.001	.001	.001	.001	.001
Talker (G)	.597	.313	.001	.001	.001	.029	.004
TG Interaction	.001	.001	.001	.146	.001	.001	.001
Experimental condition (F)	.001	.001	.001	.001	.001	.001	.001
TF Interaction	.001	.001	.001	.001	.001	.001	.001
GF Interaction	.083	.790	.043	.044	.173	.059	.004
TGF Interaction	.275	.063	.001	.076	.380	.001	.013
Age (A)	.001	.001	.001	.001	.001	.001	.001
TA Interaction	.218	.024	.357	.097	.005	.001	.010
TE Interaction	.001	.001	.001	.002	.007	.001	.001
GE Interaction	.306	.170	.012	.593	.566	.306	.552
TGE Interaction	.375	.060	.004	.651	.095	.234	.016
FE Interaction	.441	.002	.004	.166	.004	.044	.156
TFE Interaction	.583	.034	.001	.919	.800	.431	.023

TABLE 15

MEAN SCORES FOR EACH TARGET OF THE NONSENSE SYLLABLE MATERIALS BY SUBTEST

Subtest Number	Type	Voiceless Consonants								
		p	t	k	f	θ	s	ʃ	tʃ	h
1	after /a/	.77	.70	.69	.41	.23	.54	.82		
2	after /u/	.67	.62	.68	.77	.59	.49	.59		
3	after /i/	.49	.75	.74	.18	.41	.64	.81		
Mean	(average after vowel)	.64	.69	.70	.46	.41	.55	.74		
5	before /a/	.76	.73	.70	.77	.55	.76	.74	.86	.74
Mean	(average 1+2+3+5)	.67	.70	.70	.53	.45	.61	.74		

TABLE 15--Continued

Subtest Number	Type	Voiced Consonants													
		b	d	g	v	ɣ	z	m	n	ŋ	l	w	j	dʒ	r
6	before /a/	.82	.81	.77							.86	.65	.87	.84	.81
7	before /a/	.54	.80	.72	.24	.62	.72	.84	.89						
Mean	(average 6+7)	.68	.80												
4	after /a/	.70	.76	.68	.49	.12	.79	.77	.80	.85					
Mean	(average 6+7+4)	.69	.79	.72	.36	.37	.76	.89	.84						

higher overall mean scores than those with voiceless consonants.

2. There were differences in mean score between the different targets for the same consonants with different vowel environments. For example, the weak fricative consonants /f/ and /θ/ had higher intelligibility after /u/ and low intelligibility after /i/.

3. There were differences in target syllable discriminability depending on the ensemble of possible responses in a subtest. For example, the target /b/ is presented before the vowel /a/ for both subtests six and seven, yet in one subtest the mean score is 82% while in the other it is 54%.

4. Higher scores were obtained, on the average, for voiceless consonants in the initial position as opposed to the final position. This effect was greater for the fricatives than the plosives. Initial-final differences for voiced targets varied according to consonant type.

In addition to the major trends cited above, a number of interactions were observed. The largest interactions are reported below and the more detailed tabulations are presented in Appendix L.

Target Syllable-Talker Interaction. For the voiced consonants the male talker obtained higher scores than the female talker. For the voiceless targets this situation follows even though the advantage in male talker intelligibility is less.

Although for most consonants the score for the male talker was higher than those for the female (as in /d/, /ð/, /z/, /n/, /j/, /d/, /r/, /s/, /h/, and /ʃ/), there are those consonants for which the female talker obtained the higher score (/l/, /w/, /f/, /t/, /p/, and /k/). In most cases when the male obtained higher scores it was across contexts, but there were a few exceptions. For example, the target /b/ which appears in two contexts (both before and after /a/). When /ba/ is the target, as in subtests 6 and 7, the female talker obtains the higher score. In comparison when /ab/ is the target, the male talker obtains the higher score.

Target Syllable-Experimental Group Interaction. The data showing this interaction may be found in Appendix M.

A review of the mean scores for the nonsense syllables with voiced consonants reveals a number of relationships:

(a) The consonants /d/, /w/, /z/, and /ð/ are affected more by listener hearing loss than other consonants. For example, the mean score for /d/ for the conductive hearing loss group was 84%, 86.5% for the flat sensori-neural group but dropped to 67% for the severely sloping sensori-neural group of listeners.

(b) The consonants /g/, /m/, and /n/ appear to be relatively independent of listener hearing loss. The mean score for /m/

for the after /a/ vowel context was 80% for the severely sloping sensori-neural group and 78% for the conductive group of listeners.

(c) Of the consonants considered, /j/, /w/, /l/, /r/, /ŋ/, and /dʒ/ were presented in only one vowel context (after /a/), so that it is difficult to generalize these results for all consonant types. Nevertheless, these consonants showed marked differences in scores between the different experimental groups.

The mean scores for the nonsense syllables involving voiceless consonants revealed that the scores for /s/, /θ/, /f/ and /ʃ/ are relatively poorer for the listeners with the severely sloping losses, indicating that some consonants are affected relatively more by hearing loss than others.

Target Syllable-Experimental Condition Interaction. The results of the Target Syllable-Experimental Condition interaction are presented in Appendix N. As discussed earlier, the +6 dB/octave slope filter condition and the 0 dB/octave slope filter condition yielded similar scores which were significantly higher than the -6 dB/octave slope condition, the noise condition yielded by far the lowest scores.

It is indicated that:

(a) There are discrimination differences for the same consonant in the same vowel context as is shown by the consonant /b/ in

subtests 6 and 7. For example, the mean scores for subtest 6 were always higher than for subtest 7, with a difference of almost 30% in favor of subtest 6. This difference is found across all experimental conditions.

(b) Some voiced consonants are more affected by noise than others. For example, /g/, /v/, /z/, /r/, and /ð/ have lower scores in noise than /d/, /m/, /n/ and /dʒ/, but in quiet the differences are more affected by Experimental Group.

(c) In noise and in the 0 dB/octave slope filter condition it was noted that the scores for the consonants in the final position are generally lower than in the initial position context regardless of the voiced target.

(d) The advantage of the +6 dB/octave slope filter condition is most in evidence for the /ð/, /z/, /w/, and /dʒ/ consonants where the differences over the 0 dB/octave slope and -6 dB/octave slope filter conditions are in excess of ten percentage points for each consonant. It is also present for the voiceless consonants /p/, /t/, /k/, /f/, /s/, /θ/ and /ʃ/.

In addition to the main effects and interactions discussed above, it was noted that vowel context produced marked differences in the discriminability of the consonants. This is evident, for example, for the consonant /p/ which has a mean percent correct score ranging from 84.2% after /a/ to 37.8% after /i/. For many of the

consonants, scores were lowest for conditions preceding the /i/ vowel context.

Analysis of Confusion Matrices

Separate confusion matrices were generated for each subtest and Experimental Condition. In analyzing these matrices, an effort was made to separate out confusions which occurred for all Experimental Groups from confusions that differed consistently between the groups, from confusions reflecting differences in Experimental Condition, and these, in turn, from idiosyncratic confusions which occurred sporadically and were not obviously linked to any single factor.

In order to simplify the analysis, only the most common confusions were analyzed in detail. Confusions occurring at a rate of less than 10 percent were not subject to detailed analysis. Table 16 presents the most common confusions for each of the seven subtests of the nonsense syllable test, averaged across Talker, Experimental Group and Experimental Condition. An examination of the confusions indicates that the most common errors are found for the fricative to fricative substitutions, more frequently for the voiceless consonants.

The second most common group of errors occurred for the fricative to plosive errors. These are roughly evenly divided between the voiced and voiceless consonants. There were four plosive to

TABLE 16

MOST COMMON CONFUSIONS BY SUBTEST (AVERAGED ACROSS TALKER, EXPERIMENTAL GROUP AND EXPERIMENTAL CONDITION)*

Subtest Vowel Context	1 after /a/	2 after /u/	3 after /i/	4 after /a/	5 before /a/	6 before /a/	7 before /a/
	θ → f	ʃ → s	f → θ	v → b	θ → s	w → l	v → ɣ
	f → p	s → θ	θ → f	ɣ → z	ʃ → tʃ	d → g	v → b
	θ → p	p → t	p → t	ɣ → d	h → p	g → d	b → ɣ
	θ → t		s → θ	ɣ → g			

*a → b means that the /a/ is the target and /b/ is the consonant reported by the subject.

plosive errors, again roughly equally divided between the voiceless and voiced consonants. There was only one plosive to fricative error (/b/ to /ð/) and two other types (/h/ to /p/ for glottal to plosive error and /w/ to /l/ for glide to lateral error).

On the average, however, there were more voiceless than voiced confusions. This is to be expected since, as noted earlier, the percent correct identification was higher, on the average, for the voiced consonants.

Most common confusions by subtest and talker. Table 17 presents the most common confusions by subtest as well as talker (averaged over experimental group and condition). The table indicates that there are certain confusions which occur frequently for both the male and female talker: /θ/ to /f/, /f/ to /p/, /θ/ to /p/, /s/ to /θ/, /f/ to /θ/, /ð/ to /g/, /θ/ to /s/, /w/ to /l/, /v/ to /ð/, and /v/ to /b/. There are no plosive to plosive errors seen, and an almost equal number of fricative to fricative (5) and fricative to plosive (4) substitutions.

Those confusions specific to the female talker were: /ʃ/ to /s/, /k/ to /p/, /ð/ to /z/, /ð/ to /v/, /ʃ/ to /tʃ/, /d/ to /g/, /w/ to /j/ and /m/ to /n/. Note that there are no fricative to plosive errors, the most predominant errors being fricative to fricative for both voiced and unvoiced consonants.

TABLE 17

MOST COMMON CONFUSIONS BY SUBTEST AND TALKER (AVERAGED ACROSS EXPERIMENTAL GROUP AND EXPERIMENTAL CONDITION)

Subtest Vowel Context	1 after /a/	2 after /u/	3 after /i/	4 after /a/	5 before /a/	6 before /a/	7 before /a/
			Male Talker				
	θ → f f → p θ → p	p → t s → θ t → θ	f → θ p → t θ → f	v → b ð → d ɣ → g	h → p θ → s s → θ	w → l g → d w → g	v → ð b → ð v → b
			Female Talker				
	θ → f f → p θ → p	ʃ → s s → θ k → p	f → θ s → θ θ → f	ʒ → z ɣ → g ɣ → v	ʃ → tʃ θ → s θ → f	w → l d → g w → j	v → ð v → b m → n

The /p to /t/, /t/ to /θ/, /ð/ to /d/, /h/ to /p/, /g/ to /d/, /w/ to /g/, and /b/ to /ð/ confusions were all specific to the male talker. These confusions can almost be paired for the different types of confusions: plosive to plosive (/p/ to /t/ voiceless and /g/ to /d/ voiced), plosive to fricative (/t/ to /θ/ voiceless and /b/ to /ð/ voiced), /h/ to /p/ (voiceless) and /w/ to /g/ (voiced) for other errors and lastly, one fricative to plosive error (/ð/ to /d/ voiced).

Most common confusions by Subtest, Talker and Experimental Group. A further breakdown of the confusions introduced the variable of Experimental Group in addition to Subtest and Gender, and is presented in Table 18. As would be expected, there are a greater number of confusions present for the sensori-neural hearing loss groups of subjects. All the groups, however, follow the pattern already presented of more voiceless than voiced substitutions. In addition, one notes the spread of errors for the severely sloping sensori-neural group for the voiced consonants, whereas for the voiceless errors there are predominantly plosive to plosive confusions. Table 19 summarizes these findings.

Most common confusions by Subtest, Talker, Experimental Group and Experimental Condition. When the confusions were broken down to consider all four factors it was found that because of the

TABLE 18

MOST COMMON CONFUSIONS SPECIFIC TO EXPERIMENTAL GROUP, IN ADDITION TO THOSE IN TABLES 16 AND 17, FOR THE MALE AND FEMALE TALKER (AVERAGED OVER CONDITION)

		Group				
Set No.	Talker	1	2	3	4	5
1	Male	k → θ	k → f		k → p t → p	t → p
	Female				k → p	
2	Male	k → s	p → θ	θ → s p → θ	k → p t → p	k → p
	Female	p → s	t → p		t → p	θ → t
3	Male	p → k p → θ	p → θ	p → θ		f → t
	Female	p → k p → θ	k → t	p → k		k → t

TABLE 18--Continued

Set No.	Talker	Group 1	2	3	4	5
4	Male		ð → v		g → d	v → g
	Female				v → g m → n	b → d g → z v → z
5	Male		k → p	k → p f → θ	t → p	θ → f
	Female			p → k	s → θ	s → θ
6	Male	b → g w → j	l → b	w → j	b → d	l → r w → r
	Female			w → j r → l		w → r
7	Male				b → d g → d	b → d g → d
	Female	v → d	g → ð	v → b z → d	g → d z → d	z → ð g → d

TABLE 19

SUMMARY OF TYPES OF CONFUSIONS FOR EXPERIMENTAL GROUPS AND TALKERS

		Experimental Group									
		1	2	3	4	5	1	2	3	4	5
Type of Error	Gender	Voiced					Voiceless				
Plosive to plosive	M	1	0	0	4	2	1	2	1	4	3
	F	0	0	0	1	1	1	1	2	2	0
Fricative to fricative	M	0	1	0	0	0	0	0	2	2	1
	F	0	0	0	0	2	0	0	0	1	1
Fricative to plosive	M	0	0	0	0	1	0	0	0	1	1
	F	1	0	2	2	1	0	0	0	0	1
Plosive to fricative	M	1	0	0	0	0	2	3	2	0	0
	F	0	1	0	0	1	2	0	0	0	0
Other	M	1	1	1	0	2	0	0	0	0	0
	F	0	0	2	1	1	0	0	0	0	0

small sample size per target, the variability in observed confusion rates was high. Consequently it was decided to limit the analysis to only those confusions in excess of 25 percent, these being the most important confusions. The results of this tabulation are presented in Table 20. It is necessary to look at the data for the three conditions in quiet as separate from the noise condition because of the large differences in the confusion patterns for the noise condition. For example, the noise condition showed a high rate of confusions involving the nasal sounds whereas these confusions are only found in two other isolated instances for the quiet conditions, both involving the severely sloping sensori-neural groups of subjects.

In the noise situation the most frequently confused consonants, in addition to the nasals, are the /k/, /p/, /θ/, /s/, /v/, /b/ and /g/. There are some consonants for which there are errors in noise for every experimental group: /f/, /θ/, /p/, /v/, /d/, /r/, /g/, and /ɣ/. Others are present only for the high-frequency hearing losses: /z/ and /ʃ/.

In all three quiet filter conditions the number of confusions increases as the degree of high frequency sensori-neural hearing loss increases, generally involving these consonants in order of decreasing magnitude for these experimental groups: /s/, /f/, /t/, /θ/, /ʃ/, /k/, /z/, /p/, /v/, /ɣ/, /tʃ/, /dʒ/, /b/, /n/, and isolated instances of /w/, /d/, /g/, /r/ and /j/.

TABLE 20

MOST COMMON CONFUSIONS SPECIFIC TO EXPERIMENTAL CONDITION (IMPLIES NOT INCLUDED IN TABLES 16, 17 OR 18) FOR MALE AND FEMALE TALKERS

		Experimental Condition									
		OdB/Octave Filter Slope					OdB + Noise				
		Group									
Set No.	Gender	1	2	3	4	5	1	2	3	4	5
1	M	f → θ			s → p	s → f s → p k → t	f → k k → t	f → k	p → f	p → t s → t k → ʃ	s → p f → t
	F	s → θ		s → θ	s → θ	s → t s → θ	k → p f → s θ → s p → t	k → t p → t k → p θ → s	s → p	s → t s → p k → t t → k	s → p s → ʃ
2	M	k → t			s → p f → ʃ	θ → s	p → s	θ → s k → s	θ → t s → t	k → t s → p θ → s p → k	θ → k p → θ

TABLE 20--Continued

Set No.	Gender	Experimental Condition: 0dB/Octave Filter Slope, Group 1	Group 2	Group 3	Group 4	Group 5	0dB + Noise, Group 1	Group 2	Group 3	Group 4	Group 5
	F				s → k	s → k f → s	t → s	t → s	p → θ	k → θ θ → t s → t p → k	p → s ʃ → t s → t
3	M			θ → p	p → k	f → k f → p	t → θ		p → θ f → k p → f θ → t	θ → k θ → t f → t	θ → p t → θ ʃ → k s → t k → p
	F				s → p	s → p t → k θ → tʃ ʃ → θ	θ → k f → k k → t	f → k	f → k	k → t θ → t	p → k k → θ θ → t k → p t → θ s → f ʃ → θ

TABLE 20--Continued

Set No.	Gender	Experimental Condition: OdB/Octave Filter Slope, Group 1	Group 2	Group 3	Group 4	Group 5	Experimental Condition: OdB + Noise, Group 1	Group 2	Group 3	Group 4	Group 5
4	M				v → g	z → g	v → g d → n m → g b → g b → z	v → g g → d d → z m → b	g → n b → g m → ʒ m → b	v → d g → n	b → g g → ŋ
	F	g → ɣ			b → d v → g	z → g	n → m m → n v → d b → z b → d	n → m m → ŋ b → ✓ b → d d → z v → g n → ŋ	d → n ɣ → n n → m v → z v → g	g → z z → n v → n n → m ɣ → n d → g	b → g v → g
5	M				t ʃ → ʃ		k → p f → k p → h	p → h	p → h	f → p k → p t → h	k → p k → θ t → k s → ʃ

TABLE 20--Continued

Set No.	Gender	Experimental Condition: OdB/Octave Filter Slope — Group 1	2	3	4	5	Experimental Condition: OdB + Noise — Group 1	2	3	4	5
	F	p → k		p → k	p → k	ʃ → s k → p tʃ → ʃ	k → t	t → k p → t	θ → p	f → p p → k	k → s k → p
6	M				w → r	dʒ → j	l → w		l → w		l → b
	F					dʒ → j	r → l	r → l	r → j	r → l	r → j
7	M			ɣ → d			g → d	d → g g → d	m → b g → d	ɣ → d v → d ɣ → b	z → d ɣ → d v → g
	F			z → ɣ	z → ɣ	z → d	d → ɣ g → d ɣ → z m → v	ɣ → d m → ð ɣ → n d → g n → g d → ɣ	g → ɣ m → b b → v d → ɣ	g → ɣ v → z	v → z d → ɣ d → z ɣ → n

TABLE 20--Continued

Set No.	Gender	Experimental Condition: +6 dB/Octave Slope, Group 1	Group 2	Group 3	Group 4	Group 5	Experimental Condition: -6 dB/Octave Slope, Group 1	Group 2	Group 3	Group 4	Group 5
1	M				f → t		f → t t → p s → f		s → f	k → p s → p ʃ → p f → t s → θ	t → θ s → p
	F	f → θ			θ → k		s → θ ʃ → ɛ			k → t f → θ	s → f t → θ s → p ʃ → s
2	M					t → p k → s	θ → t			ʃ → θ	f → s f → θ ʃ → θ
	F	θ → s				t → p ʃ → f f → ʃ	f → ʃ				θ → p ʃ → θ

TABLE 20--Continued

Set No.	Gender	Experimental Condition: +6 dB/Octave Slope, Group 1	Group 2	Group 3	Group 4	Group 5	-6 dB/Octave Slope, Group 1	Group 2	Group 3	Group 4	Group 5
3	M				θ → p p → k		t → k			θ → p	f → k θ → t
	F				θ → p	t → k p → k p → θ	f → p	f → p		f → p θ → p	θ → p ʃ → θ f → k t → k p → k
4	M				b → d	n → ŋ				b → d	
	F									g → d v → z	n → m

TABLE 20--Continued

Set No.	Gender	Experimental Condition: +6 dB/Octave Slope, Group 1	Group 2	Group 3	Group 4	Group 5	Experimental Condition: -6 dB/Octave Slope, Group 1	Group 2	Group 3	Group 4	Group 5
5	M										t → p t → h k → t
	F	p → s p → θ				t → p				s → f t → p	t → p
6	M										
	F								r → w	r → w j → w	
7	M	b → d			z → ɣ ɣ → v					ɣ → d	
	F				v → d	d → g v → d	z → ɣ			z → ɣ ɣ → d	ɣ → z

Examination of Table 20 also reveals that the errors follow patterns already discussed, i.e., plosives are most often confused with plosives, fricatives are confused with plosives, fricatives with fricatives, and lastly fricatives with plosives. There are also the nasal confusions already mentioned and those few confusions involving the other types of consonants which occur sporadically throughout this table. It is interesting to note that while the errors occur for type of consonant, they do not involve the manner of production, i.e., voicing.

Analysis of Rating-Score Reliability Ratings

Each subject in the study rated the relative intelligibility of each of the 18 listening conditions on a scale of 1-7, the higher the rating the higher the intelligibility. In order to examine the reliability of the subjects ratings, and, in particular, whether talker gender affected the ratings, the measured intelligibility scores were plotted against the intelligibility ratings. Separate plots were obtained for the male and female talkers and for each subject for each Experimental Condition.

A second-degree polynomial was fitted to each set of data. The standard deviation of the plotted points from the fitted curve was derived for each set of data. An analysis of variance was then performed on these standard deviations to determine if the factors of

Talker Gender and Experimental Group affected variability.

The results of this analysis of variance are shown in Table 21. As is seen in this table, none of the factors was found to be statistically significant, indicating that there were no consistent effects of Talker Gender and Experimental Group on the variability of the ratings.

Once it was determined that the deviations from the fitted curves were essentially homogeneous for the two factors, the spread of the points from the mean was examined. The statistic (R^2) is the proportion of the total variation about the mean explained by the regression. Hence, the larger the R^2, the greater the association between the two variables. The R^2 values were subjected to an analysis of variance to determine whether Talker Gender or Experimental Group affected the relationship between rated and measured percent intelligibility. These results are presented in Table 22. Talker Gender was found to be not statistically significant whereas Experimental Group showed a significant effect.

The mean R^2 values for the five Experimental Groups are shown in Table 23. The data show a consistent trend in which the association between the two variables (measured and rated intelligibility) decreased systematically with increasing hearing-impairment. The largest R^2, accounting for 87.5 percent of the variance, was obtained for the normal hearing subjects, whereas for the sloping

TABLE 21

ANALYSIS OF VARIANCE ON VARIATION OF STANDARD DEVIATIONS OF PLOTTED POINTS FROM FITTED CURVES FOR INTELLIGIBILITY RANKINGS

Source Variation	Sum Squares	DF	Mean Squares	F-Ratio	Signif. Level
Gender	.00	1	.00	.43	.51
Pathology	.06	4	.01	.07	.37
Gender-pathology	.04	4	.01	.77	.54

TABLE 22

ANALYSIS OF VARIANCE ON SPREAD OF POINTS FROM THE MEAN (R^2) FOR PLOTTED AND FITTED CURVES

Source Variation	Sum Squares	DF	Mean Squares	F-Ratio	Signif. Level
Gender	.02	1	.02	.37	.549
Pathology	1.49	4	.37	6.66	.001
Gender-pathology	.03	4	.00	.15	.956

TABLE 23

MEAN R^2 VALUES FOR THE EXPERIMENTAL GROUPS

Experimental Group	Percent of Variance
Normal hearing	87.5
Conductive hearing loss	78.1
Flat sensori-neural hearing loss	67.4
Moderately sloping sensori-neural hearing loss	57.4
Severely sloping sensori-neural hearing loss	55.5

sensori-neural hearing loss subjects only 55.5 percent of the variance was accounted for by the relationship between the two variables.

CHAPTER V

DISCUSSION

The purpose of this study was to investigate the differences in speech intelligibility of sentences and nonsense syllables spoken by male and female talkers for normal hearing and hearing-impaired listeners. Of the two types of materials used, there were many (10) talkers used for the sentences, but only one talker of each gender (who also recorded the sentences) for the nonsense syllables.

One of the major findings of this study was that the sentence intelligibility scores for the female talkers were higher on the average than the scores for the male talkers. These results are consistent with those obtained by Mayer (1980) in a study on aging effects on speech perception using the same recordings.

The extent to which this finding can be generalized to male and female talkers selected at random from the population is limited. Even so, it is surprising that, contrary to the predominant complaint of the hearing impaired that the female voice is more difficult to understand, this study provided no evidence in support of this observation. Since there is no reason to expect the ten talkers considered were uniquely different from talkers in general, it is significant to

note that contrary to the expected results, no advantage was found for the male voice, and, if anything, the female talkers were more intelligible than the male talkers.

Although the five female talkers were significantly higher, on average, on the intelligibility scores, it is important to recognize that several individual male talkers were more intelligible than some of the female talkers. In addition, the differences between male and female talkers were found to vary depending upon Experimental Condition and Experimental Group.

As shown in Table 9, the degree of overlap in ranking of relative intelligibility between male and female talkers varied as a function of Experimental Group and Experimental Condition. For the most part it seemed that two of the male talkers were more intelligible than the rest, but with the female talkers one female stood out as being by far the most intelligible, two as being relatively intelligible, and two as being moderately intelligible and comparable to the least intelligible male talkers. That is, the overlap between the male and female talkers occurred at the less intelligible side of the range. It is not possible to ascertain with any degree of confidence whether the observed overlap holds for the population as a whole. Unfortunately, in order to obtain such an estimate, an immense study with a very large sample of talkers would be required.

A second consideration is that the spectrum of background

noise, i.e., Experimental Conditions, and the differences between the hearing impairments affect the measured differences between talkers. A more pertinent question might be to ask which factors affect the intelligibility of male and female talkers. The evidence suggests that the spectrum of the noise relative to the spectrum of the speech is a possibility, however, the relatively high intelligibility score for the worst condition (female voice in the presence of female noise) suggests that spectrum shape is not the only consideration.

Another consideration is that of relative levels. In this study S/N ratio was controlled which differs somewhat from real life communication in which the female voice is typically less powerful than the male voice and usually is heard at a poorer S/N ratio. The magnitude of the difference is approximately 3 dB. In a study dealing with the development of speech reception materials, Elkins and Causey (1977) reported a 10 percent increase in correct responses on CHABA sentence materials per decibel increase in presentation level for a male talker. This would indicate that the male talker could be 20-30 percent more intelligible than the female talker (in face-to-face communication) on the basis of this intensity difference alone. In the present study, when the 5-dB corrections were made for M5 (Male Talker #5), it was found that the 5-dB difference affected intelligibility by approximately 30 percent.

The male talker advantage of greater overall power should

not be present in most clinical and test environments because voice levels are typically calibrated against a VU meter. This calibration should eliminate overall level differences between male and female talkers. This helps explain why the hearing impaired patient reportedly has no greater problem understanding female, as opposed to male, audiologists in the clinic. This is especially important when considering that there are more female than male audiologists, and that 64.4% of clinicians are using live-voice techniques (Martin and Pennington, 1971), rather than the standardized male recordings available to the profession.

A matter of some concern is the way in which speech level is measured. For the present study, average peak speech levels were used to specify the speech level. This is a difficult procedure to employ because of the instrumentation problems involved (e.g., the ballistics of the meter or recording instrument used). Small differences in the specification level may occur quite easily and only a slight change in the specified speech level (e.g., 1 dB) can produce a significant change in the measured speech intelligibility scores, particularly for sentence materials where the percent intelligibility function is very steep. In this study it was assumed that any criteria effects in setting speech levels affected both male and female talkers equally.

If differences in speech power were the only differences

between male and female voices, equating speech levels should result in equal intelligibility for male and female talkers. This was not the case, however. If differences in both level and speech spectrum were the only differences of importance between male and female voices, then similar results should have been obtained for the male voice against a background of male-spectrum shaped noise as for the female voice against a background of female-spectrum shaped noise. Again, this was not the case. It is clear that more than level and spectrum differences play a part in accounting for intelligibility differences between talkers. Possible additional factors may be the consonant-vowel amplitude ratio, which has been shown by Weiss (1968) to be more closely related to intelligibility than S/N ratio. Pascoe (1974; 1975) has suggested that the differences in speech intelligibility should not be "generalized as a sex difference." Mazor, Simon, Scheinberg and Levitt (1977) suggest that there is a relationship between the frequency spectrum of female speech and relative intelligibility for hearing-impaired listeners. Harris et al. (1960) has suggested that there are certain quality factors which make some voices more intelligible than others. It is probable that the superiority of one talker, be it male or female, is determined by a combination of factors. Braida (1979) has argued for an index that would be predictive of a particular talker's intelligibility.

The results obtained in this study are consistent with the past

research dealing with inter-talker differences. Palmer (1955) for example, cites the differences between male and female voices, but found no differences between talkers used in his study, possibly because he used a very small sample of listeners with a wide age range. Kruel, Nixon, Kryter, Bell, Lang and Schubert (1968) used three talkers in their evaluation of the MRHT. Although they did not deal directly with inter-talker differences, an examination of their data reveals that the relative intelligibility of their female talker lay midway between that of their other two talkers who were male. Beyer, Webster and Dague (1969) found large talker differences in their study and were able to relate part of the list differences to one of the talkers who produced a consistently lower intelligibility score than the other talkers. House, Williams, Kuker and Kryter (1965) also found materials recorded by one talker to be consistently more intelligible than those recorded by the other talker; both of their talkers were male. Harris, Haines, Kelsey and Clack (1960) used CHABA sentences and although they were interested in the effects of low fidelity circuitry on intelligibility, they nevertheless found large inter-talker differences under identical experimental conditions.

Pascoe (1975) used one male and one female talker in his study. He amplified the level of the female voice to the level of the male talker and found that all subjects had higher scores with the male-talker recordings. Penrod (1976) studied the effects of different

talkers on the word discrimination scores of hearing-impaired listeners. His talkers were 3 males and 1 female (all trained audiologists). In this case, some of the subjects scored consistently low for an individual talker, but the talker showing the consistently low scores varied between subjects.

A factor having a pronounced effect on intelligibility is the type of hearing impairment. A sensori-neural hearing loss often affects intelligibility of speech to a much greater degree than a conductive hearing loss, particularly against a background of noise. A common finding which was observed in this study is that the speech intelligibility decreased as the degree of hearing loss in the higher frequencies increased. The data clearly showed that intelligibility scores decreased systematically in going from a flat sensori-neural hearing loss to a moderately sloping hearing loss to a severely sloping sensori-neural hearing loss. In contrast, the conductive hearing loss subjects performed relatively well and as a consequence were presented with speech at a S/N ratio equal to that used with the normal hearing subjects. On average, the intelligibility of speech in noise was not significantly different for the conductive hearing-loss group than for the normal hearing group. Another factor of importance is the age of the listener. The elderly form one of the largest patient groups in the population. They present specific problems in terms of rehabilitation because of the central

and peripheral changes which occur as part of the aging process (Mayer, 1980; Pickett, Levitt, and Bergman, 1979; Smith and Prather, 1971). The data cited in these studies are consistent with those of this study, namely that the older listeners have poorer speech discrimination than the younger listeners. In addition, there appear to be perceptual changes in the processing of consonants by the elderly. Specifically, Stevenson (1975) has suggested that there is a reduction in the number or availability of phoneme categories that the subjects were able to handle. There was the implication that consonants in initial and final positions were different in detail for the listeners, and therefore were processed differently, which may have relevance for the perception of speech among the aged. Recently, entire texts have been devoted to the issue of the geriatric patient (Schow, Christensen, Hutchinson and Nerbonne, 1978; Oyer and Oyer, 1976; Barry and Wingrove, 1977; Maurer and Rupp, 1979; 1979).

An unexpected result was that obtained for the listeners with conductive hearing losses. In this case the older subjects obtained higher scores on average than the younger listeners. On re-examination of the scores it was found that one subject of the older group obtained unusually high intelligibility scores for every listening condition, essentially setting his MCL almost to the maximum output of the equipment. This subject was unusual and his scores were so much better than those of any other listener in the study that

his data produced a reversal in the trend in which younger listeners usually obtain higher intelligibility scores than older listeners.

Since intelligibility tests are very time consuming it is of interest to observe how ratings of relative intelligibility compare to the measured intelligibility scores. Of particular interest are the ratings of relative intelligibility of male and female talkers. As shown in Table 9, there was a significant relationship between the type of hearing loss and the ability to rate measured intelligibility. The most consistent relationship between intelligibility ratings and intelligibility scores was obtained by the normal hearing group. The conductive hearing loss group showed almost the same degree of consistency; the flat sensori-neural group was poor, and the sloping sensori-neural groups were essentially the same and showed the poorest consistency. That is, the group with the poorest intelligibility scores was also the group least able to rate intelligibility consistently. Although there were slight differences in the relative ratings of intelligibility for individual talkers (e.g., M4 was rated more intelligible than M5, but M5 had higher intelligibility scores), the listeners rated the female talkers as more intelligible on the average than the male talkers. There appears to be no difference in the relative consistency of the ratings between male and female talkers.

One of the concerns was the informal observation that female talkers are not as intelligible as male talkers. Thus it was of interest

to see how the subjects rated the intelligibility of the male and female talkers used in the study. Table 24 shows the relative rankings for the experimental talkers for the five experimental conditions and groups. The structure is analagous to Table 9. The table shows similar but not identical rankings of the relative intelligibility of the various talkers. A striking difference between the two tables is that M4 was rated as more intelligible than M5 in eight of the ten cells where both M4 and M5 appear. In contrast, the intelligibility scores showed M5 to be more intelligible than M4 on average.

The main thrust of the study was the nature of the male-female talker and listener differences in the perception of speech. It was also of interest to obtain analytic data on representative male and female talkers. M5 and F5 were chosen for this purpose to record the nonsense syllables by a group of skilled phoneticians who analyzed these talkers and found them to have fluent, articulate speech. These talkers were chosen in advance of the present study. As it turned out, M5 was one of the best male talkers, but F5 was not one of the best female talkers. In comparing M5 and F5, it was found that M5 was more intelligible, and in addition, obtained higher intelligibility scores on the NST. Therefore, M5 and F5 are not directly comparable in terms of their intelligibility. M5 may be the best of the male talkers, but F5 was representative of an average level of

TABLE 24

RANKINGS OF TALKER INTELLIGIBILITY RATINGS AS ASSIGNED BY ALL LISTENERS
(WITH SIX ENTRIES PER CELL REPRESENTING THE SIX TALKERS PER CONDITION)

Noise	Talker	Experimental Group* 1	2	3	4	5
High-pass	M	M4 M3M1	M4M1M3	M1M4M3	M4M2 M3	M3M1M4
	F	F3F5F1	F3F1	F3F1 F5	F3 F5 F1	F3F1 F5
			F5			
Low-pass	M	M3 M1M2	M2M1 M3	M1M2M3	M1 M2 M3	M3 M2M1
	F	F4F3	F4	F4F3F5	F4 F3 F5	F5 F4F3
		F5	F3 F5			
White	M	M2 M1M5	M2M5M1	M5M2M1	M2M1M5	M5M1M2
	F	F2F3F1	F3F2 F1	F3F1	F2F1 F3	F3 F1
				F2		F2

*Numbers refer to group assignments used throughout the study (see Table 6).

TABLE 24--Continued

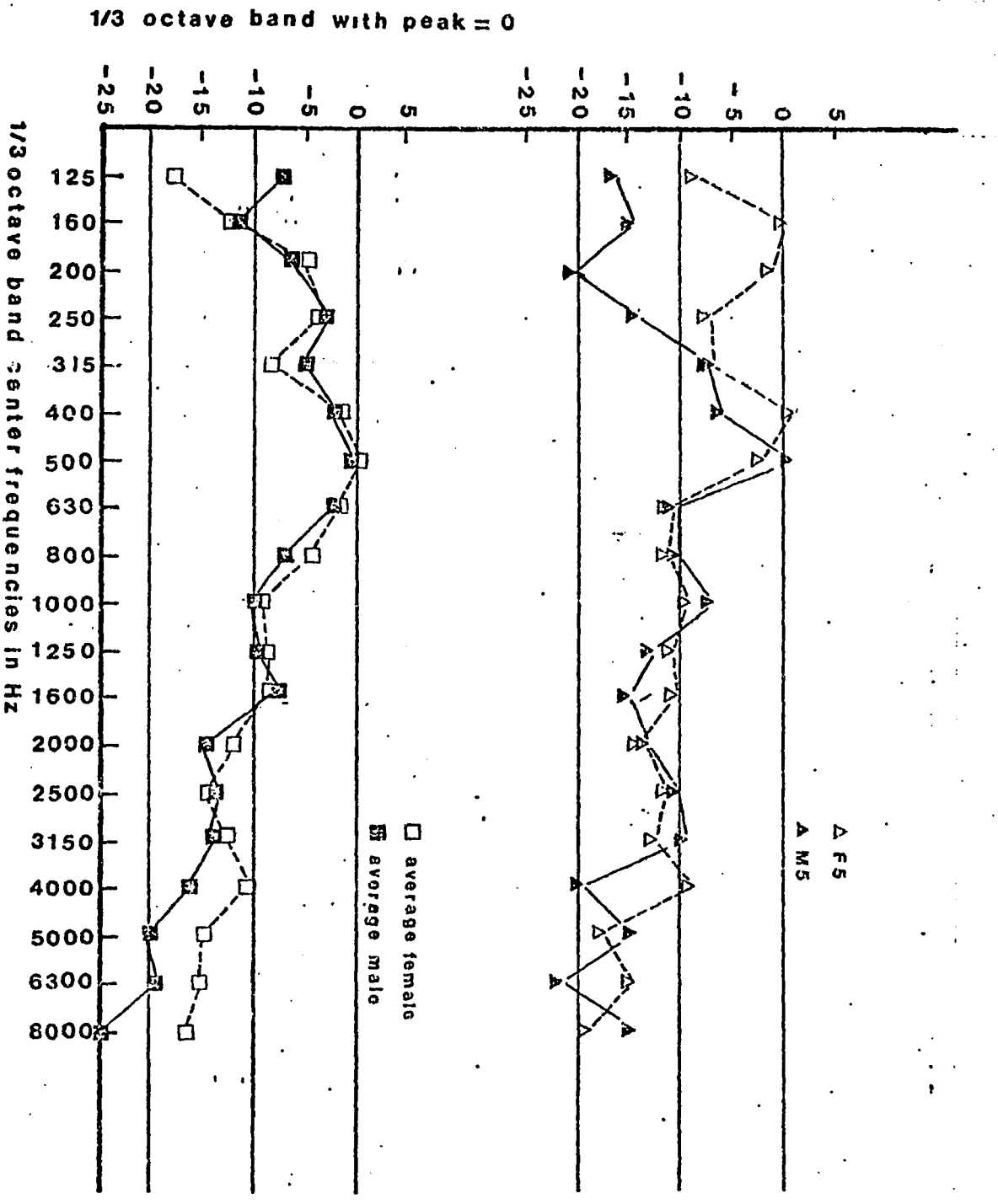
Noise	Talker	Experimental Group 1	2	3	4	5
Male-spectrum	M	M4M2M5	M2 M4M5	M2M4M5	M4M5 M2	M2M4 M5
	F	F2 F1F4	F2F4 F1	F2F4 F1	F2F1 F4	F2 F1F4
Female-spectrum	M	M3 M5M4	M5 M4 M3	M4 M5M3	M3M5M4	M4 M5 M3
	F	F4F2F5	F2 F4 F5	F2 F4 F5	F5F5F2	F4 F2 F5

intelligibility. In comparing these two talkers, the differences should be regarded as individual differences rather than male-female differences. This is revealed not only by their relative intelligibility scores, but their speech spectra also show individual differences. This is illustrated in Figure 26. The top half shows M5 and F5, the lower half shows the average 1/3 octave speech spectra for male and female talkers obtained by Pearsons, Bennett and Fidell (1977). In both cases the spectra are for speech produced in quiet. As can be seen on Pearsons et al. data, the female voice has more relative power in the high frequencies. While this average spectrum has more relative power in the higher frequencies, M5 shows about the same relative power in the high frequencies as F5. In addition, M5 was reported to articulate more clearly, and so it was not surprising that he consistently obtained higher intelligibility scores than F5.

While the male talker obtained higher intelligibility scores for all experimental groups and experimental conditions, there were some target consonants for which the female talker obtained higher scores. These consonants were /l/, /w/, /f/, /p/, /t/, /k/ (with /tʃ/ being almost the same as the male). Note that F5 consistently obtained higher discrimination than M5 for the voiceless plosives. For this talker the burst of aspiration on voiceless plosives was clearly more distinct than for M5. While there are individual

Fig. 26. Top: 1/3 octave band speech spectra of talkers M5 and F5 (who recorded nonsense syllable materials).

Bottom: 1/3 octave band average speech spectra per Pearsons et al.


1/3 octave band with peak = 0
-25
-20
-15
-10
-5
0
5
1/3 octave band center frequencies in Hz
125
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
average female
average male
F5
M5

differences between talkers, a possible hypothesis is that female talkers on average may have more audible aspiration, making voiceless plosive consonants more intelligible.

Of the four experimental conditions considered with the nonsense syllable materials, the +6 dB/octave slope condition produced the highest scores for both the male and female talker for all listeners of all ages. Previous research has shown that a high-frequency boost typically improves the intelligibility of speech (French and Steinberg, 1947). This improvement has also been observed for hearing-impaired listeners with high-frequency hearing loss (Levitt et al., 1978; Skinner, 1978). Of particular relevance is the finding of Witter and Goldstein (1971) who showed significant improvement in intelligibility for both male and female talkers with an upward frequency shift in the spectrum.

It may be noted that the nonsense syllable test is weighted toward high-frequency sounds, e.g., voiceless consonants in the final position (Levitt et al., 1978) and a positively sloping frequency response may be particularly good for these materials, but not necessarily for speech in general.

The effects of the Age and Experimental Group variables for these two talkers on the nonsense syllables are consistent with those data obtained using CHABA sentences. In general, the younger listeners obtained higher intelligibility scores than the older listeners

and the relatively flat audiometric configurations obtained higher intelligibility scores than the sloping sensori-neural groups. As before, the shape of the audiogram is an important consideration. In this case one would expect experimental groups with the moderately sloping and severely sloping sensori-neural losses to show a larger reduction score for the nonsense syllable materials since these two groups have greater high-frequency hearing losses and the test is biased towards high-frequency sounds. Note that these two groups also scored relatively poorly on the sentence materials.

Of the other two slopes considered (0 and -6 dB/octave) the scores were about the same and lower than for the +6 dB/octave slope. On the basis of audibility of the spectrum, one might have expected the 0 dB/octave slope condition to be better than the -6 dB/octave slope condition but this was not the case, and no obvious reason could be found other than that the variability of the measurements obscured any real differences.

The one noise condition used with the nonsense syllable materials was one in which the male voice was presented in the presence of the male-spectrum shaped noise and the female voice was presented in a background of female-spectrum shaped noise. In this way, all frequencies were being equally masked. These were considered the worst cases in terms of resulting intelligibility and confusions. The resulting pattern of confusions was similar to that

obtained with cafeteria noise by Dubno (1978). This is not surprising since the speech-spectrum shaped noise is similar in spectrum shape to the cafeteria noise used by Dubno, but not identical. It was found that the fricative consonants were the most difficult to identify, there were more manner confusions than in quiet (confusions in which target and response syllable were produced by a different manner, e.g., fricative and plosive), and most common place confusions (same manner) were between pairs of plosives. A comparison for the normal listeners in the two studies revealed that the /if/→ /iθ/ and /ba/→ /ða/ confusions were common in noise across experimental samples.

In general, it was noted that for the nonsense syllables plosives were most often confused with plosives, fricatives with plosives, fricatives with fricatives and lastly fricatives with plosives. It should be remembered that the response foils chosen for the nonsense syllable test were chosen so that voicing errors were not a possible alternative. The test was designed this way since voicing confusions occur least frequently with hearing-impaired persons (Levitt and Resnick, 1978; Pickett et al., 1970).

Vowel context has an important effect on the discrimination of consonants. As shown when the vowel was /a/, consonants were most easily identified; when /i/ and /u/ were the vowels, discrimination was much lower, being poorest for the /i/. For the vowel

/u/ the two formants F1 and F2 are close together, and hence the transitions are harder to perceive; for /i/ the second formant may be beyond the audible range of the hearing-impaired listener, or may be masked by noise along with F3, leaving F1 alone to be confused with the very similar and close together pair of F1 and F2 of /u/.

These results compare favorably with those obtained by Dubno (1978) and Abramovitz (1979) for both normal and hearing-impaired listeners.

Similarly, the data obtained in this study compare favorably with those of Levitt et al. (1978) where consonants in the initial position were identified more often than in the final position and that voiced consonants obtained higher intelligibility scores than unvoiced consonants.

The effect of hearing loss on the discrimination of the various speech sounds show distinct patterns. Certain consonants are more affected by hearing loss than others, e.g., /d/, /w/, /z/, / /, while in comparison, /g/, /m/, and /n/ are relatively unaffected by the hearing loss. Note that the latter group of consonants have substantial low-frequency power. Similarly, the consonant confusions show a pattern in which consonants with similar spectra are confused with each other and the greater the high-frequency contour (as in the fricatives and voiceless plosives), the greater the rate of confusion.

In summary, it is seen that there are large individual differences between talkers, and that on the average the female talkers in this study obtained higher intelligibility scores than the male talkers for the noise conditions considered with the sentence materials. However, it is extremely difficult to generalize the result to the population at large. Further, it may be more productive to look for differences relative to individual talkers rather than gender related differences. Clearly overall level and spectrum shape are important considerations in this regard, but there are presumably other variables. If these factors could be identified, the results would have important implications for the development of more effective acoustic amplifying systems for hearing-impaired people and the way in which aural rehabilitation is provided.

CHAPTER VI

SUMMARY AND CONCLUDING COMMENTS

Summary of Experiment

This study was concerned with investigating differences in intelligibility between male and female talkers for normal and hearing impaired listeners. Five subject groups of ten listeners each, composed of normal hearing listeners, listeners with conductive hearing losses, flat sensori-neural hearing losses, moderately sloping sensori-neural hearing losses and severely sloping sensori-neural hearing losses, were subdivided according to listener sex as well as older and younger groups. The subjects were presented with two types of stimulus material: the CID sentences of everyday speech (CHABA) recorded by five male and five female talkers, and recordings of a nonsense syllable test available for both a male and a female talker (who had also recorded the sentence lists).

Five types of noise were used for each talker group for the sentence materials: white noise, low-pass noise, high-pass noise, male-spectrum shaped noise and female-spectrum shaped noise. Four experimental conditions were used for the nonsense syllable materials: frequency filtering with a slope of 0 dB/octave, +6 dB/

octave and -6 dB/octave (all in quiet) and 0 dB/octave slope condition in noise at a S/N ratio of 0 dB.

Findings

The five female talkers were, on average, more intelligible than the five male talkers used in the study. However, the distribution of scores showed that some males were more intelligible than some female talkers. For the nonsense syllable materials, higher discrimination scores were obtained by the male rather than the female talker. It should be noted, however, that the male talker used for the nonsense syllables was the most intelligible talker for the sentence materials whereas the female talker generated an average intelligibility score.

Differences between experimental groups were as expected. The normal hearing listeners obtained the highest scores, followed by the conductive hearing loss group, the flat-, moderately- and severely sloping sensori-neural group.

The background noise reduced intelligibility scores in a predictable way with one exception. The speech-shaped noise was most damaging to intelligibility; high-pass and white noise were the next most damaging, with low-pass noise the least effective masker. These findings are consistent with articulation index theory.

For the nonsense syllables, the +6 dB/octave slope condition provided the highest intelligibility for both talkers for all listeners.

In addition, it was found that syllables with voiced consonants typically showed higher overall mean intelligibility scores than those with voiceless consonants; that vowel context affected intelligibility, especially for fricative consonants; and that higher scores were obtained for consonants in the initial position as opposed to the final position.

The older subjects showed consistently lower scores with one exception (the conductive group) possibly because of erratic scores obtained by one subject.

Conclusions

Male-female differences in talker intelligibility are affected by a number of complex variables and a simple generalization cannot be made that male talkers are more or less intelligible than female talkers. On average, for the talkers and noise conditions in this study, the female talkers were more intelligible than the male talkers for the sentence materials. Several important caveats should be noted. Firstly, in setting speech levels and S/N ratios, the speech signals were adjusted to be equal in power as measured by the average of the speech peaks on a level recorder. In real life, female voices are typically several dB lower than male voices. Secondly, no one method of specifying speech level is generally agreed upon, especially when talker characteristics differ in some way. Other methods of equating speech level may have yielded somewhat different results

since the slope of the performance-intelligibility function for sentence material is very steep and small differences in equating speech levels may produce large changes in the intelligibility score. Thirdly, male-female differences in intelligibility are also affected by the spectrum of the background noise. The relative S/N ratio in different frequency regions is an important factor and this varies as a function of both the speech and noise spectrum.

It is important to bear in mind that although there are average differences between male and female talkers in terms of spectrum shape, speech level and other relative acoustic parameters, there are large individual differences between talkers in terms of these parameters. For example, some males have weaker voices than some female talkers or some female talkers have a spectrum shape similar to that of a male talker, and vice versa.

Given these many differences, it is evident that there are a number of inter-acting factors, each of which may affect the intelligibility of male and female talkers differently. It may be better to look for improved methods of predicting speech intelligibility from the acoustic measurements of speech and background noise which are independent of gender.

APPENDIX A

INDIVIDUAL SUBJECT DESCRIPTIONS

Description of Subjects in Experimental Group 1
(Conductive Hearing Loss Group)

Test Ear	SRT dB HL	PB Score	MCL (HL)	MCL-SRT (SL)	Tymp.[2] Type	Re-flex	H.A.[3]	Age	Sex
Right	60	100	85	25	A_S	No	Yes	55	Male
Right	45	100	80	35	A^S	No	No	41	Female
Left	45	100	80	35	A	No	No	53	Female
Right	35	100	70	35	B	No	No	63	Female
Right	35	72	65	30	A	No	No	49	Male
Left	35	100	65	30	A	No	No	54	Female
Left	30	88	65	30	CNT[4]	CNT	No	49	Male
Left	40	96	75	35	CNT[5]	CNT	No	40	Male
Left	45	92	80	35	A	No	Yes	53	Female
Left	30	100	65	35	A	No	Yes	47	Male
Mean	40	94.8	73	32.5				50.4	

[1]Conductive hearing loss defined as at least a 30 dB air-bone gap with normal bone conduction thresholds.

[2]Type of Tympanogram based upon the classification of Jerger (1974).

[3]Refers to whether the subject had ever had any hearing aid experience prior to the study.

[4]Subject had impedance studies performed on 4/21/77 when initially seen as a patient and declined having the procedure repeated on 4/4/78 when seen as a study subject since there had been no substantial change in his hearing on the other test measures.

[5]Subject complained that his ear was very sensitive and allowed impedance tests only on the right ear where the Tympanogram was a Type C and the stapedius reflex was absent. Otologic examination on the left ear failed to discover any active pathology to account for the discomfort reported by the patient.

Description of Subjects in Experimental Group 2
(Normal Hearing Listeners)

Test Ear	SRT dB HL	PB Score	MCL (HL)	MCL-SRT (SL)	Tymp. Type	Reflex	H.A.	Age	Sex
Left	0	100	35	35	A	Yes	No	26	Female
Left	-5	100	30	35	A	Yes	No	29	Female
Right	0	100	35	35	A	Yes	No	27	Female
Left	5	100	35	30	A	Yes	No	25	Male
Right	0	96	35	35	A	Yes	No	28	Male
Right	0	100	35	35	A	Yes	No	28	Female
Left	0	100	35	35	A	Yes	No	28	Female
Right	-5	100	30	35	A	Yes	No	20	Female
Right	-5	100	30	35	A	Yes	No	27	Female
Left	0	100	35	35	A	Yes	No	24	Female
Mean	-1	99.6	33.5	34.5				26.2	

Description of Subjects in Experimental Group 3
(Flat Sensori-Neural Hearing Loss)

Test Ear	SRT dB HL	PB Score	MCL (HL)	MCL-SRT (SL)	Tymp. Type	Reflex	H. A.	Age	Sex
Left	50	100	80	30	A	Yes	Yes	37	Male
Left	30	100	65	35	A	Yes	No	43	Female
Left	40	80	75	35	A	Yes	No	47	Female
Right	45	88	75	30	A	No	Yes	48	Female
Right	35	100	65	30	A	No	No	48	Female
Left	30	80	65	35	A	Yes	No	48	Female
Left	30	84	60	30	A	Yes	No	49	Male
Right	40	92	75	35	A	Yes	No	54	Female
Right	55	76	80	25	A	Yes	Yes	68	Female
Right	35	100	65	35	A	Yes	Yes	78	Female
Mean	39	90	63	32				52	

Description of Subjects in Experimental Group 4
(Moderately Sloping Sensori-Neural Hearing Loss)

Test Ear	SRT dB HL	PB Score	MCL (HL)	SL	Tymp. Type	Re-flex	H.A.	Age	Sex
Left	40	84	75		A	Yes	No	28	Female
Right	55	80	70		A	No	Yes	30	Male
Right	35	100	70		A	Yes	Yes	39	Female
Left	30	60	65		A	Yes	No	49	Male
Left	35	96	70		A	Yes	Yes	58	Male
Left	30	80	60		A	Yes	No	65	Male
Right	55	64	85		A	No	No	66	Male
Right	45	48	70		A	No	Yes	56	Male
Left	50	68	85		A	No	Yes	68	Female
Right	30	52	60		A	Yes	Yes	70	Male
Mean	40.5	73.2	71					53.9	

Description of Subjects in Experimental Group 5
(Severely Sloping Sensori-Neural Hearing Loss)

Test Ear	SRT dB HL	PB Score	MCL	SL	Tymp. Type	Re-flex	H.A.	Age	Sex
Left	35	48	70	35	A	Yes	Yes	21	Male
Left	35	76	75	40	A	Yes	Yes	33	Female
Right	35	92	65	30	A	Yes	Yes	44	Female
Right	30	88	65	35	A	Yes	Yes	47	Male
Right	40	76	80	40	A	Yes	Yes	57	Female
Left	30	84	65	35	A	Yes	No	58	Male
Right	30	64	65	30	A	Yes	No	65	Male
Left	40	64	70	30	A	No	No	68	Male
Right	35	56	65	30	A	Yes	No	75	Female
Left	35	40	75	40	A	Yes	Yes	80	Male
Mean	34.5	68.8	63	35				54.8	

APPENDIX B

CHABA LIST ASSIGNMENTS PER TALKER

CHABA Sentence List Assignments Per Talker

CHABA List	Talker
A	Male 2
B	Female 3
C	Female 2
D	Male 4
E	Female 4
F	Male 5
G	Female 1
H	Female 5
I	Male 1
J	Male 3

APPENDIX C

MICROPHONE AMPLIFIER AND GRAPHIC LEVEL RECORDER SETTINGS FOR PREPARATION OF SPEECH-SPECTRUM MATERIALS

Settings for Equipment Used to Obtain Speech Spectrum Materials

Equipment	Control	Setting
B&K Microphone Amplifier #2603	Input Pot.	0
	Input Switch	Direct
	Meter Range	10V- 0dB- 120dB SPL
	Frequency Response Switch	External Filter
	Meter Switch	RMS Fast
	Range Multiplier	0 dB x 1
B&K BandPass Filter Set Type 1612	Input Switch	Direct
	Automatic Switching	Off
	Weighting Network	Off
	Function Selector	1/3 Octave 0 dB
B&K Graphic Level Recorder Type 2305	Input Pot.	Reference
	Input Attenuator	Set so that calibration tone is approximately 20 dB to begin, and changed as necessary
	Potentiometer Range	50
	Rectifier	RMS
	Frequency Limit	50
	Writing Speed	400/800
	Drive Shaft	1.2
	Paper Speed	1 (10)

APPENDIX D

CHABA SENTENCE TEST TAPE ASSIGNMENTS AND ENTIRE TEST PROTOCOLS PER SUBJECT

Test Protocols

Subject	NST Order 1	2	3	4	5	6	7	8	CHABA Tape	Order (CHABA)
1	6m	5m	8m	6f	8f	7m	5f	7f	1	1
2	6m	5f	8f	5m	7f	8m	7m	6f	2	2
3	7m	6f	6m	7f	5m	5f	8m	8f	3	1
4	7f	8f	6f	5m	6m	8m	5f	7m	1	2
5	6m	8m	8f	5f	7f	6f	5m	7m	2	1
6	5m	5f	6m	8f	6f	7m	8m	7f	3	2
7	8f	7f	5f	6f	5m	8m	7m	6m	1	1
8	8m	6m	7m	7f	5m	6f	5f	8f	2	2
9	8f	8m	6m	7m	7f	5m	5f	6f	3	1
10	8f	7f	5m	8m	7m	6f	6m	5f	1	2
11	5m	6f	7m	7f	6m	5f	8m	8f	2	1
12	5f	6m	7f	8f	8m	7m	5m	6f	3	2
13	6f	5f	5m	8m	7f	8f	7m	6m	1	1
14	7f	6m	5f	8f	7m	5m	8m	6f	2	2
15	6f	5m	8m	7m	8f	5f	6m	7f	3	1
16	7m	7f	8m	6m	5f	8f	6f	5m	1	2
17	5m	8m	7f	6m	7m	8f	5f	6f	]	1
18	5f	6f	8m	8f	7m	5m	7f	6m	3	2
19	6f	6m	5f	7m	8m	7f	5m	8f	1	1

Test Protocols--Continued

Subject	NST Order 1	2	3	4	5	6	7	8	CHABA Tape	Order (CHABA)
20	7m	8m	5m	7f	8f	5f	6m	6f	2	2
21	7f	7m	8f	8m	6f	5f	5m	6m	3	1
22	5m	5f	6m	6f	8f	7m	7f	8m	1	2
23	7m	5m	6m	6f	5f	7f	8m	8f	2	1
24	6f	5m	8m	7m	5f	7f	6m	8f	3	2
25	8f	6f	6m	5m	7f	5f	7m	8m	1	1
26	7m	6m	5f	8f	8m	5m	6f	7f	2	2
27	8m	7m	6f	6m	8f	5m	5f	7f	3	1
28	5f	8m	8f	6f	5m	6m	7m	7f	1	2
29	8f	7f	8m	7m	5f	6f	5m	6m	2	1
30	8f	5f	6f	5m	6m	7m	7f	8m	3	2
31	8m	5m	7m	5f	7f	8f	6m	6f	1	1
32	5m	7m	8m	6m	8f	7f	5f	6f	2	2
33	6f	5f	5m	7f	7m	8m	6m	8f	3	1
34	8m	7f	7m	5m	6f	5f	8f	6m	1	2
35	7f	7m	6f	5m	8f	5f	8m	6m	2	1
36	5m	6f	8f	5f	6m	8m	7f	7m	3	2
37	8f	8m	5f	6m	7f	7m	6f	5m	1	1
38	5f	6m	6f	7m	8m	5m	7f	8f	2	2

Test Protocols--Continued

Subject	NST Order 1	2	3	4	5	6	7	8	CHABA Tape	Order (CHABA)
39	6f	5f	8f	7m	6m	7f	5m	8m	3	1
40	5m	6f	7m	8f	5f	6m	7f	8m	1	2
41	5f	8m	7m	6m	5m	8f	6f	7f	2	1
42	6m	8f	8m	6f	7f	5f	7m	5m	3	2
43	8m	7m	5m	8f	6f	7f	6m	5f	1	1
44	8f	5f	6f	5m	7m	8m	6m	7f	2	2
45	5f	8m	8f	6f	7m	6m	7f	5m	3	1
46	6f	8m	5f	7m	6m	8f	7f	5m	1	2
47	7f	5m	6m	8f	5f	8m	7m	6f	2	1
48	7m	8f	6m	8m	6f	5m	7f	5f	3	2
49	8m	7f	7m	6f	8f	5m	6m	5f	1	1
50	7f	6m	7m	5f	6f	8f	8m	5m	2	2
51	5m	6f	7f	6m	8m	5f	7m	8f	3	1
52	8f	5m	7m	6f	7f	6m	8m	5f	1	2

	Talker	List	Noise Condition
CHABA Test Tape #1			
1.	F1	G	white noise
2.	F2	C	male shaped noise
3.	F5	H	high pass noise
4.	M3	J	high pass noise
5.	M1	I	white noise
6.	M5	F	female shaped noise
7.	M2	A	low pass noise
8.	F3	B	low pass noise
9.	M4	D	male shaped noise
10.	F4	E	female shaped noise
CHABA Test Tape #2			
1.	F3	B	high pass noise
2.	F4	E	low pass noise
3.	M1	I	low pass noise
4.	M2	A	white noise
5.	M5	F	male shaped noise
6.	M3	J	female shaped noise
7.	F2	C	white noise
8.	F1	G	male shaped noise
9.	F5	H	female shaped noise
10.	M4	D	high pass noise

	Talker	List	Noise Condition
CHABA Test Tape #3			
1.	M5	F	white noise
2.	F2	C	female shaped noise
3.	M3	J	low pass noise
4.	M2	A	male shaped noise
5.	F1	G	high pass noise
6.	M4	D	female shaped noise
7.	F5	H	low pass noise
8.	F3	B	white noise
9.	M1	I	high pass noise
10.	F4	E	male shaped noise

APPENDIX E

SUBJECT INTAKE FORM

SPEECH DISCRIMINATION STUDY
Subject Intake Form

Name: _____

Address: _____

_____

Phone: _____ Birthdate: ____ / ____ / ____
Mo. Day Year

Pure Tone Thresholds (Re: ANSI, 1969 ref. Thresh):

Air Conduction	250	500	1000	2000	4000	8000Hz
Right Ear	____	____	____	____	____	____dB
Left Ear	____	____	____	____	____	____dB

Bone Conduction					
Right Ear	____	____	____	____	____
Left Ear	____	____	____	____	____

Speech Audiometry

	SRT	PB	Level
Right	____	____	____
Left	____	____	____

Tympanometry

Type____ Right

Type____ Left

Stapedius Reflex Levels

	500	1000	2000	4000Hz
Right	____	____	____	____d
Left	____	____	____	____d

Test Ear _____

Subject Number _____

Date: _____

Hearing Aid Experience

Yes____ or No____

Ear_____ Aid_____

APPENDIX F

PRESENTATION LEVELS IN DECIBELS FOR ALL LISTENERS FOR ALL EXPERIMENTAL CONDITIONS AND TALKERS (Based Upon MCL Values for Each Subject)

Presentation Levels in Decibels (re audiometric 0) for all Listeners for all Experimental Conditions and Talkers (Based Upon MCL Values for Each Subject)

Experimental Group	Experimental Condition								
	CHABA	NST							
		0 dB/octave		+6 dB/octave		-6 dB/octave		Noise	
Talker		Male	Female	Male	Female	Male	Female	Male	Female
Normal Hearing	40 dB	40 dB	40 dB	65 dB	65 dB	55 dB	55 dB	40 dB	40 dB
	40	40	40	50	60	50	50	40	40
	35	35	35	50	55	50	50	35	35
	35	35	35	55	60	55	55	35	35
	35	35	35	60	65	55	55	35	35
	40	40	40	55	60	55	55	40	40
	35	35	35	50	50	45	45	35	35
	30	30	35	50	55	45	50	30	30
	35	35	35	60	60	55	55	35	35
	35	35	35	60	60	55	55	35	35

Appendix F--Continued

Experimental Group	Experimental Condition							
	CHABA	NST						
		0 dB/octave		+6 dB/octave		-6 dB/octave		Noise
Talker		Male	Female	Male	Female	Male	Female	Male	Female
Conductive Hearing Loss	85 dB	85 dB	85 dB	110 dB	110 dB	105 dB	110 dB	85 dB	85 dB
	80	80	80	95	100	85	85	80	80
	80	80	80	90	100	100	95	80	80
	70	70	70	75	85	70	80	70	70
	65	65	65	85	85	75	75	65	65
	65	65	65	80	80	75	80	65	65
	65	65	65	80	85	75	70	65	65
	75	75	75	95	90	85	80	75	75
	80	80	80	95	90	90	90	80	80
	65	65	65	80	90	80	80	65	65

Appendix F--Continued

Experimental Group	Experimental Condition							
	CHABA	NST						
		0 dB/octave		+6 dB/octave		-6 dB/octave	Noise	
Talker		Male	Female	Male	Female	Male	Female	Male	Female
Flat Sensori-Neural Loss	70 dB	70 dB	70 dB	90 dB	95 dB	90 dB	95 dB	70 dB	70 dB
	80	80	80	105	105	105	100	80	80
	65	65	65	80	85	80	75	65	65
	75	75	75	110	110	100	95	75	75
	65	65	65	75	80	75	85	65	65
	65	65	65	85	90	85	85	65	65
	75	75	75	90	90	80	85	75	75
	75	75	75	85	80	75	80	75	75
	80	80	80	95	95	95	90	80	75
	70	70	70	85	95	80	85	70	70

Appendix F--Continued

Experimental Group	Experimental Condition							
	CHABA	NST						
		0 dB/octave		+6 dB/octave		-6 dB/octave	Noise	
Talker		Male	Female	Male	Female	Male	Female	Male	Female
Moderately sloping Sensori-neural Hearing Loss	65 dB	65 dB	65 dB	90 dB	100 dB	80 dB	85 dB	65 dB	65 dB
	75	80	80	95	105	90	95	80	80
	75	75	75	95	95	85	85	75	75
	85	85	85	100	105	95	100	85	85
	65	65	70	85	90	80	85	70	70
	70	70	70	75	80	75	75	70	70
	65	65	65	80	85	75	75	65	65
	65	65	65	90	90	85	85	65	65
	75	75	75	90	90	75	85	75	75
	75	75	75	105	100	90	95	75	75

Appendix F--Continued

Experimental Group	Experimental Condition								
	CHABA	NST							
		0 dB/octave		+6 dB/octave		-6 dB/octave		Noise	
Talker		Male	Female	Male	Female	Male	Female	Male	Female
Severely Sloping Sensori-Neural Hearing Loss	75 dB	70 dB	75 dB	95 dB	95 dB	75 dB	75 dB	75 dB	75 dB
	65	65	65	90	100	85	90	65	65
	65	65	75	85	85	75	75	75	75
	80	80	80	95	95	85	90	80	80
	75	75	75	90	90	90	90	75	75
	65	65	65	85	95	80	80	65	65
	70	70	70	85	85	80	80	70	70
	75	75	75	100	105	85	85	75	75
	65	65	65	90	95	80	85	70	70
	70	70	70	90	95	80	85	70	70

APPENDIX G

SAMPLE OF NONSENSE SYLLABLE TEST PAGE

Page_____ Name_____

OOSH	OOK	OOS	OOP	OOF	OOT	OOTH
OOSH	OOK	OOS	OOP	OOF	OOT	OOTH
OOSH	OOK	OOS	OOP	OOF	OOT	OOTH
OOSH	OOK	OOS	OOP	OOF	OOT	OOTH
OOSH	OOK	OOS	OOP	OOF	OOT	OOTH
OOSH	OOK	OOS	OOP	OOF	OOT	OOTH
OOSH	OOK	OOS	OOP	OOF	OOT	OOTH
OOSH	OOK	OOS	OOP	OOF	OOT	OOTH

APPENDIX H

SAMPLE OF CHABA SENTENCE LIST SCORE SHEET

List G

1. I'll see you right after lunch.

2. See you later.

3. White shoes are awful to keep clean.

4. Stand there and don't move until I tell you!

5. There's a big piece of cake left over from dinner.

6. Wait for me at the corner in front of the drugstore.

7. It's no trouble at all.

8. Hurry up!

9. The morning paper didn't say anything about rain this afternoon or tonight.

10. The phone call's for you.

Subject Number _____

Date _____

Score _____

Noise Condition _____

APPENDIX I

RATING SHEETS FOR THE SENTENCE AND NONSENSE SYLLABLE MATERIALS

Intelligibility Ratings

_____	Tape 1	1 2 3 4 5 6 7
_____	Tape 2	1 2 3 4 5 6 7
_____	Tape 3	1 2 3 4 5 6 7
_____	Tape 4	1 2 3 4 5 6 7
_____	Tape 5	1 2 3 4 5 6 7
_____	Tape 6	1 2 3 4 5 6 7
_____	Tape 7	1 2 3 4 5 6 7
_____	Tape 8	1 2 3 4 5 6 7

Subject _____

Date _____

Intelligibility Ratings

_____	Talker 1	1 2 3 4 5 6 7
_____	Talker 2	1 2 3 4 5 6 7
_____	Talker 3	1 2 3 4 5 6 7
_____	Talker 4	1 2 3 4 5 6 7
_____	Talker 5	1 2 3 4 5 6 7
_____	Talker 6	1 2 3 4 5 6 7
_____	Talker 7	1 2 3 4 5 6 7
_____	Talker 8	1 2 3 4 5 6 7
_____	Talker 9	1 2 3 4 5 6 7
_____	Talker 10	1 2 3 4 5 6 7

Subject _____

Date _____

CHABA Tape _____

APPENDIX J

INTRODUCTION LETTER TO POTENTIAL SUBJECTS, CONSENT FORMS SIGNED BY SUBJECTS AND THANK YOU LETTER SENT TO SUBJECTS

Martin L. Haubenstock, President
Paul H. Abrams, 1st Vice-President
Joseph Dunn, 2nd Vice President
Robert M. Singer, Secretary
William F. Forbes, Treasurer
Louis S. Miller, Assistant Secretary-Treasurer

Harvey Schoenfeld, M.B.A., M.S., F.A.C.H.A., F.A.P.H.A., F.R.S.H.
Executive Vice President and Director
Thomas B. Doolan, M.B.A., Administrator

BMHC
75 YEARS OF CARING 1903-1978

Barnert Memorial Hospital Center

680 Broadway, Paterson, New Jersey 07514 • 201-274-8000

Hearing and Speech Department
Jacqueline L. Edelstein, M.S., CCC
Department Executive

January 15, 1978

Dear _____,

As I'm sure you know, we in the hearing health fields are always trying to find new ways and improved methods of helping those with hearing problems.

As part of this continuing process, we are currently involved in a research project dealing with some aspects of speech understanding. We are evaluating some new test materials and would appreciate your cooperation.

In the next few weeks a member of our research team will be calling you on the telephone to discuss the possibility of your participation in the study. The study involves listening to recorded samples of speech typical of those used in clinical tests. The test will take about two and a half hours of your time. The results of this may be influential upon clinical practice in Audiology.

I hope you will give this opportunity your serious consideration and will agree to participate.

Sincerely,

Maris Margulies

Maris K. Margulies, M.A., CCC
Chief Audiologist

Barnert Memorial Hospital Center
680 Broadway, Paterson, New Jersey 07514 • 201-274-8000

Martin L. Haubenstock, President
Paul H. Abrams, 1st Vice-President
Joseph Dunn, 2nd Vice President
Robert M. Singer, Secretary
William F. Forbes, Treasurer
Louis S. Miller, Assistant Secretary-Treasurer

Harvey Schoenfeld, M.B.A., M.S., F.A.C.H.A., F.A.P.H.A., F.R.S.H., Executive Vice President and Director
Thomas B. Doolan, M.B.A., Administrator

CONSENT STATEMENT
for
SPEECH DISCRIMINATION STUDY

I,_____, understand that I am to be a participant in a research study on speech discrimination which is being conducted by Maris K. Margulies at the Hearing and Speech Center of the Barnert Memorial Hospital.

I understand that there are not necessarily any direct benefits to me from my participation in the study.

I have been informed that experience in previous tests of this type have shown no risk to the listener and that the procedures conform to current clinical practice.

I understand that I may withdraw my consent to this test any time prior to or during the test and that doing so will in no way prejudicially affect my or my family's treatment at this hospital.

SIGNED_____

ADDRESS_____

_____

DATE_____

Should any additional information be desired, please contact Mrs. Margulies at (201) 274-8000, ext. 229. She will be happy to answer any inquiries which you might have.

Barnert Memorial Hospital Center
680 Broadway, Paterson, New Jersey 07514 • 201-274-8000

Martin L. Haubenstock, President
Paul H. Abrams, 1st Vice-President
Joseph Dunn, 2nd Vice President
Robert M. Singer, Secretary
William F. Forbes, Treasurer
Louis S. Miller, Assistant Secretary-Treasurer

Harvey Schoenfeld, M.B.A., M.S., F.A.C.H.A., F.A.P.H.A., F.R.S.H., Executive Vice President and Director
Thomas B. Doolan, M.B.A., Administrator

Hearing and Speech Department
Jacqueline L. Edelstein, M°S., CCC
Department Executive

Date:

Dear

I am writing to you to express my gratitude for your participation in the speech discrimination study. Your cooperation, interest and patience proved invaluable. I appreciate the hours and effort you donated to the project and are grateful for the assistance you gave in this attempt to develop improved procedures for serving individuals with impaired hearing.

Thank you again.

Sincerely,

Maris K. Margulies, M.A., CCC
Chief Audiologist

MKM/pf

APPENDIX K

ORIGINAL AND CORRECTED CHABA SENTENCE SCORES FOR TALKER M5

Original and Corrected CHABA Sentence Scores for Talker M5 in Percent Correct

Original Score (%)	Corrected Score (%)
0	29
2	32
6	39
8	41
10	44
14	50
16	53
18	55
20	56
24	60
34	70
36	72
40	75
44	78
48	81
50	82
52	83
54	85
56	86
72	93
74	94 (87 for group 2)
76	94
80	95
86	96
92	96
94	96
96	96
100	100

APPENDIX L

RESULTS OF TARGET SYLLABLE-TALKER INTERACTION ON THE CONFUSION MATRICES ON THE NONSENSE SYLLABLE TEST

Results of Target Syllable-Talker Interaction on the Nonsense Syllable Confusion Matrices

Voiced Consonants

Target	Subtest	Talker Gender: Male	Female	Mean
b	6 before /a/	.74	.83	.78
	7 before /a/	.31	.66	.48
	Mean	.53	.74	.63
	4 after /a/	.70	.65	.67
	6 + 7 + 4 Mean	.58	.71	.65
d	6 before /a/	.80	.77	.78
	7 before /a/	.83	.74	.79
	Mean	.82	.75	.79
	4 after /a/	.79	.69	.74
	6 + 7 + 4 Mean	.81	.73	.77

Voiced Consonants--Continued

Target	Subtest	Talker Gender: Male	Female	Mean
g	6 before /a/	.73	.75	.74
	7 before /a/	.70	.66	.68
	Mean	.71	.70	.71
	4 after /a/	.68	.58	.64
	6 + 7 + 4 Mean	.71	.66	.68
v	6 before /a/			
	7 before /a/	.17	.22	.19
	Mean			
	4 after /a/	.47	.43	.45
	6 + 7 + 4 Mean	.32	.32	.32
j	6 before /a/			
	7 before /a/	.62	.54	.58
	Mean			
	4 after /a/	.13	.06	.10
	6 + 7 + 4 Mean	.37	.30	.38

Voiced Consonants--Continued

Target	Subtest	Talker Gender: Male	Female	Mean
z	6 before /a/			
	7 before /a/	.80	.52	.66
	Mean			
	4 after /a/	.75	.75	.75
	6 + 7 + 4 Mean	.77	.64	.71
m	6 before /a/			
	7 before /a/	.87	.77	.82
	Mean			
	4 after /a/	.75	.76	.76
	Mean	.81	.76	.79
n	6 before /a/			
	7 before /a/	.92	.86	.89
	Mean			
	4 after /a/	.84	.73	.79
	Mean	.88	.79	.83

Voiced Consonants--Continued

Target	Subtest	Talker Gender		
		Male	Female	Mean
ɹ	4 after /a/	.85	.83	.84
l	6 before /a/	.84	.86	.85
w	6 before /a/	.52	.64	.58
j	6 before /a/	.87	.84	.86
dʒ	6 before /a/	.88	.75	.81
r	6 before /a/	.89	.66	.77

Appendix L--Voiceless Consonants

Target	Subtest	Talker Gender: Male	Female	Mean
p	1 after /a/	.74	.77	.75
	2 after /u/	.53	.74	.64
	3 after /i/	.34	.48	.41
	Mean	.54	.66	.60
	5 before /a/	.80	.68	.74
	1+2+3+5 Mean	.60	.67	.64
t	1 after /a/	.56	.70	.63
	2 after /u/	.62	.60	.61
	3 after /i/	.74	.65	.69
	Mean	.64	.65	.64
	5 before /a/	.71	.64	.68
	1+2+3+5 Mean	.66	.65	.65

Voiceless Consonants--Continued

Target	Subtest	Talker Gender: Male	Female	Mean
k	1 after /a/	.66	.66	.66
	2 after /u/	.54	.67	.61
	3 after /i/	.81	.57	.69
	Mean	.67	.63	.65
	5 before /a/	.62	.67	.65
	1+2+3+5 Mean	.66	.64	.65
f	1 after /a/	.34	.39	.37
	2 after /u/	.71	.42	.72
	3 after /i/	.14	.17	.16
	Mean	.40	.43	.41
	5 before /a/	.72	.76	.74
	1+2+3+5 Mean	.48	.51	.50

Voiceless Consonants--Continued

Target	Subtest	Talker Gender: Male	Female	Mean
θ	1 after /a/	.21	.17	.19
	2 after /u/	.53	.58	.52
	3 after /i/	.32	.41	.37
	Mean	.35	.37	.36
	5 before /a/	.59	.46	.52
	1+2+3+5 Mean	.41	.39	.40
s	1 after /a/	.50	.35	.43
	2 after /u/	.46	.38	.42
	3 after /i/	.60	.49	.55
	Mean	.52	.41	.46
	5 before /a/	.78	.62	.70
	1+2+3+5 Mean	.58	.46	.52

Voiceless Consonants--Continued

Target	Subtest	Talker Gender: Male	Female	Mean
ʃ	1 after /a/	.82	.75	.79
	2 after /u/	.55	.40	.48
	3 after /i/	.82	.73	.78
	Mean	.73	.70	.68
	5 before /a/	.85	.54	.70
	1+2+3+5 Mean	.76	.61	.69
tʃ	5 before /a/	.84	.85	.84
h	5 before /a/	.73	.72	.72

APPENDIX M

RESULTS OF TARGET SYLLABLE-EXPERIMENTAL GROUP INTERACTION ON THE CONFUSION MATRICES FOR THE NONSENSE SYLLABLE TEST

Results of Target Syllable-Experimental Group Interaction on the Nonsense Syllable Confusion Matrices

Voiced Consonants

Target	Subtest	Experimental Groups[1] #2	#1	#3	#4	#5	Mean
b	6 before /a/	.91	.78	.86	.70	.78	.78
	7 before /a/	.73	.51	.53	.49	.38	.48
	Mean	.82	.64	.69	.60	.58	.63
	4 after /a/	.81	.70	.79	.59	.58	.66
	6+7+4 Mean	.82	.66	.72	.59	.58	.64
d	6 before /a/	.86	.90	.90	.62	.65	.77
	7 before /a/	.81	.84	.82	.81	.67	.79
	Mean	.84	.87	.86	.72	.66	.78
	4 after /a/	.79	.75	.82	.67	.72	.74
	6+7+4 Mean	.82	.83	.85	.70	.68	.76

[1]Experimental Group numbers refer to those used throughout the study:
1. conductive hearing loss
2. normal hearing
3. flat sensori-neural loss
4. moderately sloping sensori-neural hearing loss
5. severely sloping sensori-neural hearing loss.

Voiced Consonants--Continued

Target	Subtest	Experimental Groups #2	#1	#3	#4	#5	Mean
g	6 before /a/	.85	.75	.81	.63	.73	.73
	7 before /a/	.84	.73	.80	.55	.58	.67
	Mean	.85	.74	.81	.59	.66	.70
	4 after /a/	.83	.66	.78	.47	.60	.63
	6+7+4 Mean	.84	.71	.80	.55	.64	.68
v	6 before /a/						
	7 before /a/	.44	.16	.16	.18	.24	.19
	Mean						
	4 after /a/	.63	.50	.47	.39	.42	.44
	6+7+4 Mean	.53	.33	.32	.27	.33	.32
ɣ	6 before /a/						
	7 before /a/	.77	.72	.66	.44	.46	.57
	Mean						
	4 after /a/	.27	.13	.10	.08	.05	.09
	6+7+4 Mean	.52	.43	.38	.26	.26	.33

Voiced Consonants--Continued

Target	Subtest	Experimental Groups #2	#1	#3	#4	#5	Mean
z	6 before /a/ 7 before /a/ Mean	.90	.79	.72	.51	.62	.66
	4 after /a/	.91	.88	.84	.64	.58	.73
	6+7+4 Mean	.91	.83	.78	.58	.60	.70
m	6 before /a/ 7 before /a/ Mean	.89	.84	.85	.71	.84	.81
	4 after /a/	.82	.77	.77	.66	.80	.75
	6+7+4 Mean	.86	.81	.81	.68	.82	.78
n	6 before /a/ 7 before /a/ Mean	.89	.91	.89	.85	.89	.88
	4 after /a/	.82	.80	.82	.73	.78	.78
	6+7+4 Mean	.86	.85	.86	.79	.84	.83

Voiced Consonants--Continued

Target	Subtest	Experimental Groups #2	#1	#3	#4	#5	Mean
ŋ	4 after /a/	.89	.84	.84	.87	.78	.83
l	6 before /a/	.89	.90	.84	.82	.81	.84
w	6 before /a/	.87	.72	.72	.35	.51	.58
j	6 before /a/	.92	.89	.90	.75	.85	.85
dʒ	6 before /a/	.93	.85	.86	.79	.72	.81
r	6 before /a/	.88	.85	.77	.67	.81	.78

Appendix M--Voiceless Consonants

Target	Subtest	Experimental Groups #2	#1	#3	#4	#5	Mean
p	1 after /a/	.83	.82	.83	.71	.61	.74
	2 after /u/	.76	.60	.72	.64	.57	.63
	3 after /i/	.76	.41	.52	.41	.29	.41
	Mean	.78	.61	.69	.58	.49	.59
	5 before /a/	.83	.70	.79	.73	.74	.74
	1+2+3+5 Mean	.79	.63	.72	.62	.55	.63
t	1 after /a/	.90	.79	.77	.51	.41	.62
	2 after /u/	.68	.71	.60	.55	.56	.60
	3 after /i/	.91	.74	.77	.68	.55	.69
	Mean	.83	.75	.72	.58	.50	.64
	5 before /a/	.91	.80	.83	.59	.41	.66
	1+2+3+5 Mean	.85	.76	.74	.58	.48	.64

Voiceless Consonants--Continued

Target	Subtest	Experimental Groups #2	#1	#3	#4	#5	Mean
k	1 after /a/	.79	.66	.82	.50	.61	.65
	2 after /u/	.90	.72	.72	.54	.42	.60
	3 after /i/	.87	.72	.84	.67	.50	.68
	Mean	.83	.70	.79	.57	.51	.64
	5 before /a/	.86	.70	.73	.62	.52	.64
	1+2+3+5 Mean	.86	.70	.77	.58	.52	.64
f	1 after /a/	.58	.50	.47	.23	.27	.37
	2 after /u/	.92	.77	.86	.73	.47	.71
	3 after /i/	.31	.12	.26	.15	.09	.16
	Mean	.60	.46	.53	.37	.28	.41
	5 before /a/	.86	.77	.84	.69	.62	.73
	1+2+3+5 Mean	.67	.54	.61	.45	.36	.49

Voiceless Consonants--Continued

Target	Subtest	Experimental Groups #2	#1	#3	#4	#5	Mean
θ	1 after /a/	.40	.27	.18	.21	.11	.19
	2 after /u/	.82	.60	.66	.46	.36	.52
	3 after /i/	.59	.45	.43	.23	.34	.36
	Mean	.61	.44	.42	.30	.27	.36
	5 before /a/	.66	.66	.64	.43	.34	.52
	1+2+3+5 Mean	.62	.49	.48	.33	.29	.40
s	1 after /a/	.90	.66	.52	.52	.37	.42
	2 after /u/	.74	.55	.65	.25	.23	.42
	3 after /i/	.92	.83	.57	.39	.35	.46
	Mean	.85	.68	.58	.56	.28	.53
	5 before /a/	.94	.85	.78	.56	.56	.69
	1+2+3+5 Mean	.87	.72	.63	.56	.35	.57

Voiceless Consonants--Continued

Target	Subtest	Experimental Groups #2	#1	#3	#4	#5	Mean
ʃ	1 after /a/	.92	.88	.91	.76	.51	.77
	2 after /u/	.92	.71	.64	.35	.21	.48
	3 after /i/	.91	.83	.88	.80	.57	.77
	Mean	.92	.81	.81	.64	.43	.67
	5 before /a/	.84	.74	.83	.69	.53	.70
	1+2+3+5 Mean	.90	.77	.81	.65	.45	.68
tʃ	5 before /a/	.92	.87	.91	.85	.70	.83
h	5 before /a/	.82	.81	.75	.63	.66	.71

APPENDIX N

RESULTS OF TARGET SYLLABLE-EXPERIMENTAL CONDITION INTERACTION ON THE CONFUSION MATRICES FOR THE NONSENSE SYLLABLE TEST

Results of Target Syllable-Experimental Condition Interaction on the Nonsense Syllable Confusion Matrices

Voiced Consonants

Target	Subtest	Experimental Condition 0 dB	+6 dB	-6 dB	Noise	Mean
b	6 before /a/	.81	.87	.81	.61	.77
	7 before /a/	.52	.58	.52	.30	.48
	Mean	.67	.72	.66	.45	.61
	4 after /a/	.80	.80	.80	.23	.66
	6+7+4 Mean	.71	.75	.71	.38	.64
d	6 before /a/	.81	.86	.79	.67	.78
	7 before /a/	.87	.85	.82	.57	.78
	Mean	.84	.86	.81	.62	.78
	4 after /a/	.81	.88	.84	.35	.72
	6+7+4 Mean	.83	.86	.82	.53	.76

Voiced Consonants--Continued

Target	Subtest	Experimental Condition 0 dB	+6 dB	-6 dB	Noise	Mean
g	6 before /a/	.87	.90	.84	.23	.71
	7 before /a/	.72	.88	.78	.26	.66
	Mean	.79	.89	.81	.24	.68
	4 after /a/	.71	.79	.69	.32	.63
	6+7+4 Mean	.77	.86	.77	.27	.67
v	7 before /a/	.13	.18	.24	.15	.19
	4 after /a/	.46	.60	.56	.18	.45
	7+4 Mean	.35	.43	.36	.18	.33
ɣ	7 before /a/	.60	.74	.64	.30	.57
	4 after /a/	.09	.13	.09	.05	.09
	7+4 Mean	.35	.43	.36	.18	.33
z	7 before /a/	.71	.82	.58	.51	.66
	4 after /a/	.72	.90	.75	.58	.74
	7+4 Mean	.71	.86	.67	.55	.70

Voiced Consonants--Continued

Target	Subtest	Experimental Condition				
		0 dB	+6 dB	-6 dB	Noise	Mean
m	7 before /a/	.89	.89	.81	.63	.80
	4 after /a/	.89	.85	.82	.35	.73
	7+4 Mean	.89	.87	.82	.49	.77
n	7 before /a/	.90	.90	.89	.86	.89
	4 after /a/	.85	.89	.76	.61	.78
	Mean	.87	.89	.82	.73	.83
ŋ	4 after /a/	.91	.91	.85	.62	.82
l	6 before /a/	.86	.93	.89	.65	.83
w	6 before /a/	.62	.73	.53	.43	.58
j	6 before /a/	.89	.91	.80	.77	.85

Voiced Consonants--Continued

Target	Subtest	Experimental Condition				
		0 dB	+6 dB	-6 dB	Noise	Mean
dʒ	6 before /a/	.79	.89	.80	.74	.81
r	6 before /a/	.86	.91	.79	.49	.76

Appendix N--Voiceless Consonants

Target	Subtest	Experimental Condition 0 dB	+6 dB	-6 dB	Noise	Mean
p	1 after /a/	.84	.84	.83	.43	.74
	2 after /u/	.76	.78	.79	.18	.56
	3 after /i/	.37	.52	.51	.24	.41
	Mean	.66	.71	.71	.27	.57
	5 before /a/	.76	.87	.90	.35	.72
	1+2+3+5 Mean	.68	.75	.76	.30	.61
t	1 after /a/	.69	.81	.51	.47	.62
	2 after /u/	.68	.69	.71	.32	.60
	3 after /i/	.68	.78	.73	.57	.69
	Mean	.69	.76	.65	.45	.64
	5 before /a/	.75	.84	.59	.47	.66
	1+2+3+5 Mean	.70	.78	.64	.46	.64

Voiceless Consonants--Continued

Target	Subtest	Experimental Condition 0 dB	+6 dB	-6 dB	Noise	Mean
k	1 after /a/	.71	.82	.73	.32	.65
	2 after /u/	.65	.74	.60	.42	.60
	3 after /i/	.72	.82	.76	.44	.68
	Mean	.69	.79	.70	.39	.59
	5 before /a/	.75	.80	.80	.20	.65
	1+2+3+5 Mean	.71	.80	.72	.34	.64
f	1 after /a/	.39	.51	.40	.17	.37
	2 after /u/	.69	.83	.64	.69	.71
	3 after /i/	.19	.18	.13	.11	.15
	Mean	.42	.51	.39	.39	.41
	5 before /a/	.84	.85	.81	.38	.72
	1+2+3+5 Mean	.53	.59	.49	.38	.49

Voiceless Consonants--Continued

Target	Subtest	Experimental Condition				
		0 dB	+6 dB	-6 dB	Noise	Mean
θ	1 after /a/	.20	.23	.17	.16	.19
	2 after /u/	.65	.74	.51	.19	.52
	3 after /i/	.25	.61	.31	.28	.36
	Mean	.37	.52	.33	.26	.37
	5 before /a/	.57	.59	.62	.31	.52
	1+2+3+5 Mean	.42	.54	.40	.27	.41
s	1 after /a/	.39	.67	.31	.32	.42
	2 after /u/	.44	.58	.44	.22	.42
	3 after /i/	.49	.75	.40	.52	.54
	Mean	.44	.67	.38	.35	.46
	5 before /a/	.71	.82	.64	.61	.69
	1+2+3+5 Mean	.51	.71	.45	.42	.52

Voiceless Consonants--Continued

Target	Subtest	Experimental Condition 0 dB	+6 dB	-6 dB	Noise	Mean
ʃ	1 after /a/	.77	.87	.69	.80	.78
	2 after /u/	.37	.70	.34	.47	.47
	3 after /i/	.75	.88	.70	.75	.77
	Mean	.63	.82	.58	.67	.67
	5 before /a/	.76	.78	.79	.44	.69
	1+2+3+5 Mean	.66	.81	.63	.61	.68
tʃ	5 before /a/	.84	.91	.83	.75	.83
h	5 before /a/	.82	.88	.84	.26	.70

BIBLIOGRAPHY

Abramovitz, R. S. The effects of multichannel compression amplification and frequency shaping on speech intelligibility for hearing-impaired subjects. Ph.D. dissertation, City University of New York, 1979.

American Speech-Hearing-Language Association, personal communication, 1975.

Barry, J. R. and Wingrove, C. R., eds. Let's learn about aging. New York: John Wiley and Sons, 1977.

Bell, D. W. and Kreul, E. J. Reliability of the modified rhyme test for hearing. J. Speech Hear. Res., 1972, 15, 287-295.

Benson, R. W. and Hirsh, I. J. Some variables in audio-spectrometry. J. Acoust. Soc. Amer., 1953, 25, 499-505.

Beyer, M. R., Webster, J. C., and Dague, D. M. Re-valudation of the clinical test version of the modified rhyme words. J. Speech Hear. Res., 1969, 12, 374-378.

Black, R. and Levitt, H. Evaluation of shape gain handset for telephone users with impaired hearing. Unpublished Bell Telephone Laboratories Memorandum, 1969.

Briada, L. Personal communication, 1979.

Brandy, W. T. Reliability of voice tests of speech discrimination. J. Speech Hear. Res., 1966, 9, 461-465.

Brownlee, V. A. Statistical theory and methodology in science and engineering. 2 Ed. New York: Wiley, 1965.

Bryant, W. S. A phonographic acoumeter. Archives of Otology, 1904, 33, 438-443.

Campbell, G. A. Telephonic intelligibility. Philosophy Magazine and Journal of Science, 1910, 19 (109), 152-159.

Carhart, R. Monitored live-voice as a test of auditory acuity. J. Acoust. Soc. Amer., 1946, 17, 338-349.

_____. Problems in the measurement of speech discrimination. Archives of Otolaryngol., 1965, 82, 253-260.

_____ and Tillman, T. W. Interaction of competing speech signals with hearing losses. Arch. Otolaryngol., 1970, 91 (3), 273-279.

_____, Tillman, T. W., and Johnson, K. R. Binaural masking of speech by periodically modulated noise. J. Acoust. Soc. Am., 1966, 39, 1037.

Cooper, J. C. and Cutts, B. P. Speech discrimination in noise. J. Speech Hear. Res., 1971, 14, 332-337.

Davis, H. and Silverman, S. R. (eds.), Hearing and Deafness (4th ed.), New York: Holt, Rinehart and Winston, 1978.

Dirks, D. D. and Bower, D. Masking effects of speech competing messages. J. Speech Hear. Res., 1969, 12, 229-245.

_____ and Wilson, R. The effect of spatially separated sound sources on speech intelligibility. J. Speech Hear. Res., 1969, 12, 5-38.

_____, Wilson, R. H., and Bower, D. R. Effect of pulsed masking on selected speech materials. J. Acoust. Soc. Am., 1969, 46, 898-906.

Dubno, J. Predicting consonant confusions in noise on the basis of acoustical analysis. Ph.D. dissertation, City University of New York, 1978.

Dunn, H. K. and White, S. D. Statistical measurements on conversational speech. J. Acoust. Soc. Am., 1940, 11, 278-288.

Egan, J. P. Articulation testing methods I. OSRD Report No. 383, 1942.

_____. Articulation testing methods II. OSRD Report No. 3802, 1944.

_____. Articulation testing methods. Laryngoscope, 1948, 58, 955-991.

Eldert, E. and Davis, H. The articulation function of patients with conductive deafness. Laryngoscope, 1951, 61, 896-909.

Elkins, E. and Causey, G. D. Development of the speech reception test. DHEW publication No. (HRA) 78-1345, series 2, number 71, U.S. Department of Health Education and Welfare, 1977.

Fairbanks, G. Test of phonemic differentiation: the rhyme test. J. Acoust. Soc. Am., 1958, 30, 596-599.

Fletcher, H. Speech and hearing. Princeton: Van Nostrand, 1929.

_____ and Steinberg, J. C. Articulation testing methods. Bell System Technical Journal, 1929, 8, 806-854.

_____. Articulation testing methods. J. Acoust. Soc. Am., 1929, 1, 1-97.

French, N. R. and Steinberg, J. C. Factors governing the intelligibility of speech sounds. J. Acoust. Soc. Am., 1947, 19, 90-119.

Gardner, M. B. Effect of noise, system gain, and assigned task on talking levels in loudspeaker communication. J. Acoust. Soc. Am., 1966, 40 (5), 955-965.

Gengel, R. W. and Marks, M. A male-female speaker interaction for speech discrimination in noise with elderly compared to college age listeners. Paper presented at the meeting of the Acoustical Society of America, 1980.

Goetzinger, C. P. Word discrimination testing. In J. Katz (ed.), Handbook of clinical audiology. Baltimore: Williams and Wilkins, 1972.

Grubb, P. A phonemic analysis of half-list speech discrimination tests. J. Speech Hear. Res., 1963, 6, 271-275.

_____. Some considerations in the use of half-list speech discrimination tests. J. Speech Hear. Res., 1963, 6, 294-297.

Harris, J. D., Haines, H. L., Kelsey, P. A., and Clack, T. D. The relationship between speech intelligibility and the electroacoustic characteristics of low fidelity circuitry. J. Auditory Res., 1960, 1, 357-381.

Hirsch, I. J. Clinical application of the two Harvard auditory tests. J. Speech Hear. Dis., 1947, 12, 151-158.

_____. The measurement of hearing. New York: McGraw Hill, 1952.

_____, Davis, H., Silverman, S. R., Reynolds, E., Eldert, E., and Benson, R. W. Development of materials for speech audiometry. J. Speech Hear. Dis., 1952, 17, 321-337.

_____, Reynolds, E. G., and Joseph, M. Intelligibility of different speech materials. J. Acoust. Soc. Am., 1954, 26, 530-538.

House, A. S., Williams, C. E., Hecker, M. H. L., and Kryter, K. D. Articulation testing methods: consonantal differentiation with a closed response set. J. Acoust. Soc. Am., 1965, 37, 158-166.

Hudgins, C. V., Hawkins, J. E., Karlin, J. E., and Stevens, S. S. The development of recorded auditory tests for measuring hearing loss for speech. Laryngoscope, 1947, 57, 52-89.

Jerger, J. Clinical experience with impedance audiometry. Arch. Otolaryngol., 1970, 92, 311-324.

Johnson, C. and Young, L. The masking effectiveness of speech versus speech modulated filtered noise. Paper presented at the meeting of the American Speech-Hearing-Language Association, 7 November 1974.

Kacena, S. and Nicholls, S. Differential masking of numbers embedded in continuous discourse. Paper presented at the meeting of the American Speech & Hearing Association, 7 November 1974.

Kalikow, D. N., Stevens, K. N., and Elliot, L. L. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J. Acoust. Soc. Am., 1977, 61, 1337-1351.

Korn, T. S. Effect of psychological feedback on conversational noise reduction in rooms. J. Acoust. Soc. Am., 1954, 26, 793-794.

Kreul, E. J., Bell, D. W., and Nixon, J. C. Factors affecting speech discrimination test difficulty. J. Speech Hear. Res., 1969, 12, 281-287.

_____, Nixon, J. C., Kreyter, K. D., Bell, D. W., Lang, J. S. and Schubert, E. D. A proposed test of speech discrimination. J. Speech Hear. Res., 1968, 11, 536-552.

Kryter, K. D. The effects of noise on man. J. Speech Hear. Dis., Supp. No. 1, 1950.

_____. Masking and speech communication in noise. In The effects of noise on man. New York: Academic Press, 1970.

Levitt, H. Methods of evaluating hearing aids. Scand. Audiol. Supp. 1978, 6, C. Ludvigsen and J. Barford (eds.).

_____, Collins, M. J., Dubno, J., Resnick, S. B., and White, R. E. C. Development of a protocol for the prescriptive fitting of a wearable master hearing aid. Communication Sciences Laboratory Report #11, City University of New York Graduate School, New York, 1978.

_____, and Resnick, S. B. Speech reception by the hearing-impaired: methods of testing and the development of new tests. In Sensorineural hearing impairment and hearing aids (eds. C. Ludvigsen and J. Barfod), Scand. Audiol. Supp., 1978, 6, 107-130.

Martin, F. N. and Pennington, C. D. Current trends in audiometric practice. ASHA, 1971, 13, 671-677.

Maurer, J. F. and Rupp, R. R. Hearing and aging: a guide to rehabilitation. New York: Grune and Stratton, 1979.

_____ and Rupp, R. R. Hearing and aging: tactics for intervention. New York: Grune and Stratton, 1979.

Mayer, C. Personal communication, 1980.

Mazor, M., Simon, H., Scheinberg, J., and Levitt, H. Moderate frequency compression for the moderately hearing impaired. Communication Sciences Laboratory Report No. 9, City University of New York Graduate School, New York, 1976.

_____, Simon, H., Scheinberg, J., and Levitt, H. Moderate frequency compression for the moderately hearing impaired. J. Acoust. Soc. Am., 1977, 62, 1273-1278.

Medical Research Council: Hearing aids and audiometers. Report of the Committee on Electro-Acoustics. London, His Majesty's Stationery Office, Special Report Series No. 261, 1947.

Miller, G. A. The masking of speech. Psychological Bulletin, 1947, 44, 105-129.

_____ and Garner, W. R. The masking of tones by repeated bursts of noise. J. Acoust. Soc. Am., 1948, 20, 691-696.

_____, Heise, G. A., and Lichten, W. The intelligibility of speech as a function of the context of the test materials. J. Exp. Psych., 1951, 41, 329-335.

_____ and Licklider, J. C. R. The intelligibility of interrupted speech. J. Acoust. Soc. Am., 1950, 22, 167-173.

Oyer, H. J. and Oyer, E. J. Aging and communication. Baltimore: University Park Press, 1976.

Palmer, J. M. The effects of speaker differences on the intelligibility of phonetically balanced word lists. J. Speech Hear. Dis., 1955, 20, 192-195.

Palva, T. Studies of hearing for puretones and speech in noise. Acta Oto-laryngol., 1955, 45, 231-243.

Pascoe, D. P. Frequency responses of hearing aids and their effects on the speech perception of eight hard-of-hearing subjects. Ph.D. dissertation, Washington University, 1974.

_____. Frequency responses of hearing aids and their effects on the speech perception of hearing-impaired subjects. Annals of Otol. Rhinol. Laryngol.: Suppl. 23, vol. 86, 1975.

Pearsons, K. S., Bennett, R. L., and Fidell, S. Speech levels in various noise environments. U.S. Environmental Protection Agency, Environmental Health Effects Research Series, EPA-600/1-77-025, 1977.

Penrod, J. P. The effects of different speakers on the word discrimination scores of persons with sensori-neural hearing impairment. Ph.D. dissertation, Florida State University, 1976.

Pickett, J. M. Limits of direct speech communication in noise. J. Acoust. Soc. Am., 1958, 30, 278-281.

_____, Bergman, M., and Levitt, H. Aging and speech understanding. In Sensory systems and communication in the elderly, vol. 10, eds. J. M. Ordy and K. Brizzee. New York: Raven Press, 1979.

_____, Martin, E. S., Johnson, D., Smith, S. B., Daniel, Z., Willis, D., and Otis, W. On patterns of speech feature reception by deaf listeners. In Fant, G. (ed.), Proc. Int. Symp. on Speech Communication and Profound Deafness. Stockholm, A. G. Bell Assoc., Washington, D.C., 1970, 119-133.

Pollack, I. Masking by a periodically interrupted noise. J. Acoust. Soc. Am., 1955, 27, 353-355.

Resnick, S. B., Dubno, J. R., Hoffnung, S., and Levitt, H. Phoneme errors on a nonsense syllable test. J. Acoust. Soc. Am., 1975, 58, S114 (A).

Ross, M., Huntington, D., Newby, H., and Dixon, R. Speech discrimination of hearing impaired listeners in noise. J. Aud. Res., 1965, 5, 47-72.

Schow, R. L., Christensen, J. M., Hutchinson, J. M., and Nerbonne, M. A. Communication disorders of the aged. Baltimore: University Park Press, 1978.

Shapiro, M. T., Melnick, W., and VerMeulen, V. Effects of modulated noise on speech intelligibility of people with sensorineural hearing loss. Annals Otol. Rhinol. Laryngol., 1972, 81, 241-248.

Silverman, S. R. and Hirsh, I. J. Problems related to the use of speech in clinical audiometry. Annals Otol. Rhinol. Laryngol., 1955, 64, 1234-1244.

Skinner, M. W. Speech intelligibility in noise-induced hearing loss: effects of high frequency compensation. Ph.D. dissertation, Washington University, 1976.

Smith, R. A. and Prather, W. F. Phoneme discrimination in older persons under varying signal-to-noise conditions. J. Speech Hear. Res., 1971, 14, 630-638.

Speaks, C. and Jerger, J. Method for measurement of speech identification. J. Speech Hear. Res., 1965, 8, 185-194.

Stevenson, P. W. Responses to speech audiometry and phonemic discrimination patterns in the elderly. Audiology, 1975, 14, 185-231.

Tillman, T. W. and Olsen, W. O. Speech audiometry. In J. Jerger (ed.), Modern developments in audiology (2nd ed.), New York: Academic Press, 1973.

Traul, G. N. and Black, J. W. The effect of content on aural perception of words. J. Speech Hear. Res., 1963, 8, 363-369.

Weiss, M. S. A study of the relation between consonant-vowel amplitude ratio and talker intelligibility. M.S. thesis, Perdue University, 1968.

Wilson, R. H. and Carhart, R. Influence of pulsed masking on the threshold of spondees. J. Acoust. Soc. Am., 1969, 46, 998-1010.

Witter, H. L. and Goldstein, D. P. Quality judgments of hearing aid transduced speech. J. Speech Hear. Res., 1971, 14, 312-322.