12-1981

Magnitude Estimation of the Quality and Intelligibility of Degraded Speech

Theresa Smith
Western Michigan University

Follow this and additional works at: http://scholarworks.wmich.edu/masters_theses
Part of the Speech and Hearing Science Commons, and the Speech Pathology and Audiology Commons

Recommended Citation
http://scholarworks.wmich.edu/masters_theses/1798

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master's Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
MAGNITUDE ESTIMATION OF THE QUALITY AND INTELLIGIBILITY OF DEGRADED SPEECH

by

Theresa Smith

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Speech Pathology and Audiology

Western Michigan University
Kalamazoo, Michigan
December 1981
The purpose of the study was to identify within-group and between-group differences in perceived quality and perceived intelligibility of degraded speech for listeners with normal and impaired hearing. Following a visual magnitude estimation task, 10 subjects with normal hearing and 10 subjects with sensory-neural hearing impairment listened to connected speech samples degraded by seven levels of harmonic distortion and estimated the magnitude of their quality and intelligibility. Log average quality estimates and log average intelligibility estimates varied linearly with log degradation values for each group. The slopes of the log-log functions were interpreted as measures of perceptual sensitivity. Slopes for each listening task varied considerably within each listener group. Differences among the slopes were not statistically significant as a function of groups, tasks, or interaction of groups and tasks. The results were compared with those of related studies and implications for future research were presented.
ACKNOWLEDGMENTS

Thank you to the members of my thesis committee: Gary Lawson, Hal Bate, and John Hanley. A special debt is owed Gary Lawson, my thesis advisor, for unselfishly sharing his ideas, materials, and time. Special thanks also to John Hanley for his guidance with the manuscript and to Bill Dawson for his technical assistance.

Appreciation is extended to the subjects who volunteered for the study, to Alvin Davis of the Constance Brown Hearing and Speech Center who assisted in the acquisition of subjects, and to Alver Rongstad who was the talker for the instructions.

Personal thanks to my husband, Jim, who knows the many reasons, and to Bill and Kohnna Winfield for their support and encouragement.

Theresa Smith
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. If copyrighted materials were deleted you will find a target note listing the pages 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.
SMITH, THERESA ANNE
MAGNITUDE ESTIMATION OF THE QUALITY AND
INTELLIGIBILITY OF DEGRADED SPEECH.

WESTERN MICHIGAN UNIVERSITY, M.A., 1981
**TABLE OF CONTENTS**

ACKNOWLEDGMENTS .............................................. ii
LIST OF TABLES .............................................. v
LIST OF FIGURES .............................................. vii

**CHAPTER**

I  BACKGROUND AND PURPOSE .............................. 1
   Introduction ...................................... 1
   Statement of the Problem ......................... 11

II  METHOD AND PROCEDURES ........................ 14
   Subjects .......................................... 14
   Stimuli .................................. 16
   Procedures ........................................ 20

III  RESULTS ............................................. 25
   Introduction ...................................... 25
   Data Reduction .................................... 25
   Statistical Procedures .......................... 27
   Analysis Procedures for Magnitude Estimates ....... 28
   Analysis Procedures for Slopes ..................... 39
   Reliability ........................................ 40

IV  DISCUSSION ......................................... 48
   Introduction ...................................... 48
   Log Geometric Mean Magnitude Estimates .......... 49
   Slopes ........................................... 50
   Reliability ....................................... 51

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
# Table of Contents—Continued

**CHAPTER**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implications</td>
<td>53</td>
</tr>
<tr>
<td>V SUMMARY AND CONCLUSIONS</td>
<td>58</td>
</tr>
<tr>
<td>Purpose</td>
<td>58</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>59</td>
</tr>
<tr>
<td>Findings</td>
<td>61</td>
</tr>
<tr>
<td>Conclusions</td>
<td>62</td>
</tr>
</tbody>
</table>

**APPENDICES**

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Calibration Waveforms Used to Obtain Desired Percentages of Total Harmonic Distortion</td>
<td>64</td>
</tr>
<tr>
<td>B Transcript of Six Stimulus Passages</td>
<td>65</td>
</tr>
<tr>
<td>C Crossbreak Matrix: Presentation Orders for Passages and Random Orders for Comparison Degradation Levels</td>
<td>66</td>
</tr>
<tr>
<td>D Written Instructions and Sample Response Sheet</td>
<td>67</td>
</tr>
<tr>
<td>E Informed Consent Release Form</td>
<td>71</td>
</tr>
<tr>
<td>F Audiological Screening Form</td>
<td>72</td>
</tr>
</tbody>
</table>

**BIBLIOGRAPHY**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>73</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. A Summary of Ages, Two-Frequency Average Thresholds, Test Ear Discrimination Scores, and the Means and Standard Deviations for All Subjects .................................. 17

2. Pure Tone Thresholds (dB) and Median Thresholds as a Function of Ear (R, L) and Frequency (Hz) for the Hearing Impaired Subjects ........................................... 18

3. Mean Slopes, Standard Deviations, and Ranges for Groups 1 (Normal) and 2 (Impaired) as a Function of Listening Tasks ....................................................... 31

4. Summary of Analysis of Variance in Log Geometric Mean SQMEs Across Trials for Group 1 as a Function of Log Percent Undegraded by Linear Rectification ............. 34

5. Summary of Analysis of Variance in Log Geometric Mean SIMEs Across Trials for Group 1 as a Function of Log Percent Undegraded by Linear Rectification ............. 35

6. Summary of Analysis of Variance in Log Geometric Mean SQMEs Across Trials for Group 2 as a Function of Log Percent Undegraded by Linear Rectification ............. 35

7. Summary of Analysis of Variance in Log Geometric Mean SIMEs Across Trials for Group 2 as a Function of Log Percent Undegraded by Linear Rectification ............. 36

8. Results of Test for Linear Trend in Log Geometric Mean SQMEs for Group 1 as a Function of Log Percent (%) Undegraded by Linear Rectification ......................... 37

9. Results of Test for Linear Trend in Log Geometric Mean SIMEs for Group 1 as a Function of Log Percent (%) Undegraded by Linear Rectification ......................... 37

10. Results of Test for Linear Trend in Log Geometric Mean SQMEs for Group 2 as a Function of Log Percent (%) Undegraded by Linear Rectification ....................... 38

11. Results of Test for Linear Trend in Log Geometric Mean SIMEs for Group 2 as a Function of Log Percent (%) Undegraded by Linear Rectification ....................... 38

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
List of Tables—Continued

12. Approximate Percentages of Variance in Log Geometric Mean Magnitude Estimates Due to Log Degradation Levels That Could Be Accounted for by a Linear Equation ................................................ 39

13. Results of a Two-Way Analysis of Variance in Slopes as a Function of Two Listener Groups (Group 1 and Group 2) and Two Listening Tasks (Quality and Intelligibility) ........................................ 41

14. Pearson Product-Moment Correlation Coefficients (r) Between the Visual Magnitude Estimates of Circle Size (in²) for Each Subject's Trial 2 and Trial 3 Stimuli ................................................ 42

15. Within-Subject Correlation Coefficients (Pearson r) Between Seven Trial 1 and Trial 2 Magnitude Estimates From Each Task for Each of 20 Subjects ........ 44
LIST OF FIGURES

1. Frequency Response Curve for the Grason-Stadler 162 Speech Audiometer with TDH-49 Earphones .................. 23

2. Mean Log Geometric Mean Modulus-Free SQMEs for Groups 1 and 2 Plotted as a Function of Log Percent (%) Undegraded by Linear Rectification ................. 29

3. Mean Log Geometric Mean Modulus-Free SIMEs for Groups 1 and 2 Plotted as a Function of Log Percent (%) Undegraded by Linear Rectification .................. 30

4. Mean Slopes for Groups Plotted as a Function of Listening Tasks ........................................ 32

5. Mean Slopes for Listening Tasks Plotted as a Function of Groups ........................................ 33

6. "Average" Within-Subject Test-Retest Correlation Coefficients Between Seven Trial 1 and Trial 2 SQMEs and SIMEs for Two Groups of 10 Subjects .................. 45

7. Within-Group Correlation Coefficients Between Trial 1 and Trial 2 Slopes for Each Group Under Each Condition .................. 46
CHAPTER I

BACKGROUND AND PURPOSE

Introduction

Speech quality judgments represent a potential clinical tool in the field of audiology, particularly in the area of hearing aid evaluations. Traditional hearing aid evaluation procedures are generally based on the work of Carhart (Millin, 1980). They continue to rely on monosyllabic word discrimination tests (Martin & Forbis, 1978; Martin & Pennington, 1971) which were derived from materials originally designed for evaluating speech transmission systems with trained listeners (French & Steinberg, 1947). Several reports have indicated that such tests do not effectively or reliably rank hearing aids (Berger, 1978; Chial & Hayes, 1974; Jerger, Speaks, & Malmquist, 1966). Shore, Bilger, and Hirsh (1960) found that discrimination test scores, obtained from PB lists, may not reliably differentiate among hearing aids. Another major flaw attributed to word discrimination tests is that they do not accurately predict listener performance with connected speech (Chial & Hayes, 1974; Jerger et al., 1966). Despite the apparent limitations of traditional word tests, there has been a minimum of research related to alternative or supplemental clinical methods for hearing aid evaluation.
Quality

Speech quality judgments are frequently used clinically on an informal basis. Although based upon a paucity of research, they may be a valid alternative for estimation of hearing aid effectiveness. The majority of studies in this area have employed a paired-comparison paradigm.

Jeffers (1960) obtained quality judgments for amplified speech from 32 subjects with conductive hearing losses. She recorded a 1-minute passage of running speech through five hearing aids which she categorized from good to poor based on manufacturers' specifications of electroacoustic characteristics. Low gain aids were set for 50 dB gain at 1000 Hz, the high gain aids for 60 dB. The recordings were arranged in pairs and played to each subject through a sound-field speaker at 60 dB re: normal speech reception threshold. The subjects then indicated which passage of the pair they preferred, based on the quality of the speech. Results demonstrated that the subjects consistently preferred the passage reproduced through the aid with the better electroacoustic characteristics.

Witter and Goldstein (1971) investigated the relationship between the electroacoustic properties of hearing aids and quality judgments of 30 normal hearing listeners. A speech passage of 20-second duration was tape recorded by a male and a female speaker through five different hearing aids, representing a range of electroacoustic characteristics. Subjects listened to pairs of passages presented at 50 dB sensation level re: threshold at 1000 Hz and
decided which one of the pair "sounded best." The results, examined separately for male and female voice stimuli, revealed that the preference judgments correlated strongly with transient response, frequency range, and harmonic distortion measures.

Weldele and Millin (1975) studied the relationship between intelligibility scores and listener judgments of quality. Ten subjects with sensory-neural hearing loss gave preference judgments on pairings of four hearing aids. Subjects wore the aids, coupled to their own earmolds, as they listened to recorded speech presented in sound field at 40 dB HL. The gain of each aid was adjusted for comfort and understanding by the subjects. Following the listening task, aided SRT's and aided speech discrimination scores were measured with each aid. Subjects then judged which aid of the pair sounded "better." The results strongly supported the conclusion that there is a positive relationship between aided discrimination scores and listener judgment of perceived quality.

Punch and Ciechanowski (1977) obtained quality judgments from normal hearing and hearing-impaired listeners for both speech and musical stimuli. The speech material consisted of a 30-second passage recorded by both a male and a female talker. The speech and music were recorded through five hearing aids with their volumes set 12 dB below SSPL. The subjects, 10 normal hearing listeners and 10 listeners with sensory-neural hearing loss, listened to 10 recorded pairings of each stimulus. For normal hearing listeners the level of presentation was 66 dB SPL while the hearing-impaired listeners received the stimuli at their own MCL's. Subjects chose the passage of
each pair which they preferred on the basis of sound quality. Results indicated that quality preferences were quite similar between groups and in rank order within each group.

Yonovitz, Bickford, Lozar, and Ferrell (1978) investigated the effect of electroacoustic distortions on quality judgments of speech and music by normal hearing listeners. The stimuli were recorded through 12 different hearing aids and presented to 20 subjects at 65 dB SPL through earphones. Subjects judged the dissimilarity between the pairs and indicated their preference within each pair. The performance judgments were ranked for both speech and music and were correlated with electroacoustic distortion measurements. Findings revealed that hearing aids with better electroacoustic characteristics were preferred and that harmonic distortion was one of the features influencing speech quality judgments.

Major disadvantages of the paired-comparison paradigm are that each hearing aid must be paired with every other aid and that the result, theoretically, is an ordinal scale. Recently, a few studies (e.g., Barry & Kidd, 1981; Harris & Goldstein, 1980; Lawson, 1980) have used magnitude estimation as an alternative. Magnitude estimation is a procedure by which the quality of signals may be assessed singly to produce a ratio scale. The subject is presented with a standard stimulus by which he/she judges the value of subsequent stimuli. The judgments represent psychological responses to physical stimuli and, according to Stevens (1957), represent a power function (i.e., $\psi = K\phi^B$) in which equal stimulus ratios produce equal sensation ratios. When logarithms are taken, the power function becomes a
linear equation:

\[ \log \psi = \beta \log \phi + \log K \]

where \( \psi \) represents a psychological magnitude, \( \phi \) represents the stimulus magnitude, and \( K \) is a constant. This formula describes a straight line in log-log coordinates and the exponent \( \beta \) becomes the slope of the line (Stevens, 1975). The slope varies with different physical stimuli which produce different subjective responses and therefore can be used to compare perceptual events.

In the magnitude estimation procedure, the numerical value of the standard stimulus may be assigned by the experimenter (fixed modulus design) or by the subject (free modulus design). Stevens (1957) recommended the free modulus design because he obtained the most satisfactory power function for loudness with this method and because he preferred it to the constraint of the fixed modulus design.

Harris and Goldstein (1980) investigated the feasibility of using magnitude estimation to make hearing aid quality judgments. Ten subjects with sensory-neural hearing impairment were evaluated in a sound suite and in a reverberant room while wearing four different hearing aids. Speech discrimination scores were obtained using recorded W-22 word lists presented at 50 dB HL in quiet and in the presence of speech noise (+10 signal to noise ratio). For the quality judgment task, the stimulus was a speech passage recorded by a male talker and presented at 50 dB HTL. The gain of each hearing aid was adjusted to MCL for each subject. In a free modulus design, subjects assigned numbers to each of the four hearing aids based on the quality of the speech passage. Retests were conducted after an
interval of at least 24 hours. Results demonstrated that a clear basis for making a hearing aid recommendation was not provided by speech discrimination testing for any of the subjects while magnitude estimation provided a basis in eight out of 10 subjects. The results suggest that magnitude estimation procedures are sensitive to differences among hearing aids which are not revealed through discrimination tests. Further research is needed to determine the nature of these differences.

Lawson (1980) obtained speech quality magnitude estimates from 12 normal hearing and 12 sensory-neurally impaired hearing subjects using speech stimuli degraded by low-pass filtering, high-pass filtering, and linear rectification. All subjects underwent visual magnitude estimation training and screening prior to the listening task. The speech stimuli consisted of 10-second passages recorded by a female talker. Pairs of passages, a standard and a comparison, were presented over earphones at 40 dB re: the two-frequency pure tone average threshold. The comparison represented one of seven degrees of degradation within each type. A free modulus design was employed in which each subject assigned an arbitrary number to the quality of the standard in each set and estimated the quality of the comparison relative to the standard. Results indicated that a power function existed between the magnitude estimates and degradation levels of each of the three degradation types. Normal hearing subjects were more sensitive to linear rectification than to low-pass filtering and high-pass filtering, while hearing-impaired subjects were less sensitive to low-pass filtering than to high-pass filtering and linear
rectification. These findings imply that predictable differences exist between perceptual judgments of normal hearing individuals and hearing-impaired individuals and that these differences could have clinical significance related to hearing aid selection procedures.

Barry and Kidd (1981) employed both magnitude estimation and magnitude production procedures to develop a psychometric function for physical distortion of speech. The speech stimulus consisted of a single sentence presented at an overall level of 80 dB HL to 10 young adults with normal hearing. The signal was routed through a two-channel speech audiometer so that the input to one ear was the undistorted standard sentence while the input to the other ear was the comparison sentence which varied in second order harmonic distortion between 10 and 100%. The standard sentence was assigned a modulus of 100 in both procedures. For the magnitude estimation procedure, subjects assigned a number to the comparison that represented how distorted it sounded compared to the standard. For the magnitude production procedure, subjects adjusted the comparison to correspond to arbitrary values of 33, 50, 75, 150, 200, and 300. Half the subjects performed the estimation task first and half performed the production task first. The results of the two methods were combined into a single log magnitude-log degradation function to eliminate procedural biases. The slope of this function (.83) indicated that perceived distortion grows more slowly than physical distortion. The results suggest a linear relationship between speech perception and harmonic distortion but do not specifically define the perceptual attribute involved. Research in which different perceptual attributes
are clearly specified would make it possible to compare the slopes of the resulting functions.

In summary, speech quality studies employing a paired-comparison paradigm have yielded somewhat contradictory findings regarding discrimination tests and quality judgments. The inconclusive results reflect the need for improved research methods. Recent studies (Barry & Kidd, 1981; Harris & Goldstein, 1980; Lawson, 1980) have demonstrated that magnitude estimation may be a viable technique for this purpose and that further research in this area is warranted.

**Intelligibility**

While research on quality judgments has provided relevant data, the fact remains that intelligibility is a key factor underlying hearing aid evaluations. The relationship between the intelligibility of amplified speech and the electroacoustic characteristics of hearing aids has been investigated in several studies using noise, distorted speech, and a variety of speech stimuli.

Discrimination scores for PB word lists were derived for four degrees of distortion by Zerlin and Burnett (1960). Response effectiveness was determined for recorded PB word lists with 0, 30, 50, and 100% second order harmonic distortion for subjects with normal hearing, conductive losses, and sensory-neural hearing losses. Findings showed that subjects with normal hearing and conductive losses were less affected by distortion than subjects with sensory-neural impairment.
Zerlin (1962) asked 21 listeners with sensory-neural hearing loss to judge the intelligibility of speech samples recorded through six different hearing aids set at 50 dB of gain. He taped 30-second passages of speech in the presence of cafeteria noise with a signal-to-noise ratio of +5 dB. The subjects listened to the recordings in pairs presented at 40 dB re: spondee threshold and rated their intelligibility. Following the comparison test, speech discrimination scores were obtained on the subjects using half-lists of CID W-22 recordings which had been recorded in the same manner as the speech samples. Results indicated that the speech discrimination scores did not differentiate among the hearing aids while the performance scores yielded definite discriminations among five of the six aids. Retests demonstrated that the paired comparison judgments were fairly reliable.

Harris, Haines, Kelsey, and Clack (1961) related measurements of electroacoustic characteristics of hearing aids to speech intelligibility. The stimuli used were 300 CID sentences recorded by five different speakers (ranging in age from 8 to 68 years) under three different conditions (natural speech, denasalized speech, and interrupted speech). The sentences were recorded through seven different hearing aids in a combination of 10 different conditions. Discrimination scores were obtained for 20 normal hearing and 20 hearing impaired subjects who listened to the speech at their individual MCL's. In a second experiment 10 additional conditions were added involving controlled amounts of harmonic distortion and intermodulation distortion. In a third experiment 19 separate aid-receiver combinations
were used. The authors concluded that harmonic distortion was the main factor in degradation of speech intelligibility. However, they reported that speech was not appreciably degraded until harmonic distortion exceeded 20%.

Jerger et al. (1966) evaluated the word discrimination scores of six normal hearing listeners and six listeners with sensory-neural hearing loss, using tape recordings of PAL-8 lists in the presence of competing speech (signal-to-noise ratio of -6 dB for hearing impaired subjects, -12 dB for normal hearing subjects). The tapes were recorded through three hearing aids, with average harmonic distortion levels of 4, 10.9, and 16%, and presented to the normal hearing subjects at 64 dB SPL and to the hearing impaired subjects at the individual level which produced maximum performance on a preliminary test. All subjects completed six trials for each aid. The performance data for the majority of subjects ranked the aids in the same order, which was inversely related to the percentage of harmonic distortion.

In a similar investigation, Jirsa and Hodgson (1970) studied the effects of harmonic distortion on three groups of subjects, 12 with normal hearing, 12 with conductive losses, and 12 with sensory-neural losses. The speech intelligibility tests which they used included PAL-8 Harvard Sentence Test, CID Auditory Test W-22, and Northwestern University Auditory Test Number 6 presented in quiet and with a competing signal (signal-to-noise ratio of -4 dB). All tests were recorded through three hearing aids with harmonic distortion levels of 5.7, 20, and 29.7% and presented to the subjects at 40 dB sensation level. These measures differentiated among hearing aids at levels of
harmonic distortion above 20% which was in agreement with the find-
ings of Harris et al. (1961).

Bode and Kasten (1971) used a closed set consonant identifica-
tion task to examine the effect of harmonic distortion on speech in-
telligibility. The Modified Rhyme Test was mixed with white noise
(signal-to-noise ratio of 0 dB) and recorded through a body-type
hearing aid. Five conditions were employed in which harmonic distor-
tion values were less than 1, 5, 15, 25, and 35%. Thirty-four normal
hearing subjects were randomly assigned to three of the five test
conditions presented at 70 dB SPL. Results demonstrated that scores
decreased 15 to 29% as a function of increased distortion.

Research relating speech intelligibility to electroacoustic
characteristics of hearing aids has generally revealed that intelligi-
bility decreases as electroacoustic characteristics become poorer.
Harmonic distortion was identified as the major factor in degradation
of speech intelligibility (Harris et al., 1961; Jerger et al., 1966;
Jirsa & Hodgson, 1970) and an inverse relationship between the two
was found (Harris et al., 1961; Jerger et al., 1966). However, har-
monic distortion did not generally degrade performance until it ex-
ceeded 20%.

Statement of the Problem

Research on quality judgments of speech by normal hearing and
hearing impaired listeners has suggested a direct relationship be-
tween quality and the electroacoustic characteristics of the amplifi-
cation system. A similar relationship has been observed in studies
relating electroacoustic characteristics to speech intelligibility. Unfortunately, the relationship between quality and intelligibility is poorly understood. According to Licklider (1946) amplitude distortion affects quality more severely than it does intelligibility. Staab (1972) reported that peak clipping greatly reduces sound quality but doesn't significantly decrease speech discrimination ability. Although there is interest in the use of speech quality judgments as a clinical tool, it has not been demonstrated that quality judgments reliably reflect the intelligibility of speech nor has the perceptual relationship between the two measures been established. Magnitude estimation is a procedure by which these perceptual events can be examined and compared.

This study examined magnitude estimates of both speech quality and speech intelligibility by 10 normal hearing listeners and 10 sensory-neurally impaired listeners as a function of seven degrees of signal degradation. The log geometric means of two within-session speech quality magnitude estimates and two within-session speech intelligibility magnitude estimates were plotted as a function of log degree of degradation for the two listener groups. The log-log functions were examined to answer the following questions:

1. Is there a statistically significant trend for the log geometric mean magnitude estimates for each task (quality and intelligibility) to be influenced by change in degree of degradation for each listener group?

2. If a statistically significant trend is present for either task, what is the lowest order equation required to provide a satisfactory (i.e., statistically significant) fit to the data for each listener group?
3. If a linear trend is present for both tasks, is there a statistically significant difference among the slopes of the log-log functions as a function of: (a) listener group, (b) listening task, or (c) interaction of group and task?
Subjects

Subjects for the study were 20 adult volunteers, 10 with normal hearing (Group 1) and 10 with sensory-neural hearing impairment (Group 2) who met the criteria for the study.

Listeners With Normal Hearing

The subjects in Group 1 demonstrated normal hearing by:

(a) passing a pure tone screening test at hearing levels of 15 dB (re: ANSI S3.6-1969) for 250, 2000, 4000, and 6000 Hz;

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

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

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

(e) sustaining stable acoustic reflexes for 10 seconds at 500 and 1000 Hz (Anderson, Barr, & Wedenberg, 1970);

(f) reporting no history of otologic surgery, vertigo, or hearing loss; and
(g) obtaining a speech discrimination score of 90% or better in the test ear on a commercial version of the Northwestern University Auditory Test Number 6 (NU Auditory Test No. 6) presented at 40 dB above the two-frequency average threshold (Fletcher, 1950).

Listeners With Sensory-neural Hearing Loss

The subjects in Group 2 were hearing aid users with previously diagnosed bilateral sensory-neural hearing loss who met the following criteria:

(a) hearing threshold levels greater than 25 dB at two or more test frequencies (250, 500, 1000, 2000, 4000, and 6000 Hz) as measured by pure tone air conduction tests (Carhart & Jerger, 1959);

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

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

(d) speech discrimination score in the ear with the better two-frequency average threshold within the 60 to 90% range on the NU Auditory Test No. 6 presented at 40 dB re: two-frequency average threshold (Fletcher, 1950).

Criteria for the hearing-impaired subjects could not be more narrowly defined due to the limited population that was available.

Description of Subjects

Group 1 subjects ranged in age from 23 to 37 years with a mean of 31.10 years; Group 2 subjects ranged in age from 21 to 48 years.
with a mean of 34.30 years. For Group 1 the average thresholds for 500 and 1000 Hz ranged from 0 to 5 dB HTL (re: ANSI S3.6-1969) for the test ear with a mean of 1.60 dB; for Group 2 the two-frequency average threshold (Fletcher, 1950) ranged from 5 to 56 dB for the test ear with a mean of 39.20 dB. All of the subjects in Group 2 were hearing aid users. Length of hearing aid use ranged from 3 months to 23 years with a mean of 13.38 years. Ages, two-frequency average thresholds, test ear discrimination scores, and means and standard deviations for all subjects are summarized in Table 1. Pure tone thresholds and median thresholds as a function of ear and frequency for the hearing impaired subjects are shown in Table 2.

Stimuli

Visual Magnitude Estimation Stimuli

A visual magnitude estimation training and screening program, developed by Lawson (1980) was used to teach the subjects how to perform a magnitude estimation task and to assess the consistency of the subjects' performance. The stimuli consisted of three sets of seven pairs of geometric forms (circles and squares) presented on slides by a rear screen projector.

Speech Stimuli

The speech stimuli consisted of recorded passages representing seven different degrees of degradation by harmonic distortion. The harmonic distortion was produced by adjusting the output of a variable
Table 1
A Summary of Ages, Two-Frequency Average Thresholds, Test Ear Discrimination Scores, and the Means and Standard Deviations for All Subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age Yrs.</th>
<th>Two-Frequency Average Threshold</th>
<th>% Discrimination Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R dB</td>
<td>L dB</td>
</tr>
<tr>
<td>Group 1 (Normal)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>34</td>
<td>0a</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>8</td>
<td>0a</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>3a</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>3</td>
<td>0a</td>
</tr>
<tr>
<td>5</td>
<td>37</td>
<td>0a</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>5a</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>5a</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>32</td>
<td>0a</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>28</td>
<td>3a</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>37</td>
<td>0a</td>
<td>0</td>
</tr>
<tr>
<td>X</td>
<td>31.10</td>
<td>2.70</td>
<td>3.80</td>
</tr>
<tr>
<td>SD</td>
<td>4.53</td>
<td>2.75</td>
<td>3.71</td>
</tr>
</tbody>
</table>

Group 2 (Impaired)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>41</td>
<td>5b</td>
<td>10</td>
<td>88</td>
</tr>
<tr>
<td>12</td>
<td>48</td>
<td>15</td>
<td>10b</td>
<td>74</td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>35b</td>
<td>43</td>
<td>84</td>
</tr>
<tr>
<td>14</td>
<td>36</td>
<td>55</td>
<td>37b</td>
<td>78</td>
</tr>
<tr>
<td>15</td>
<td>36</td>
<td>55</td>
<td>48b</td>
<td>84</td>
</tr>
<tr>
<td>16</td>
<td>31</td>
<td>48b</td>
<td>58</td>
<td>84</td>
</tr>
<tr>
<td>17</td>
<td>33</td>
<td>63</td>
<td>55b</td>
<td>88</td>
</tr>
<tr>
<td>18</td>
<td>26</td>
<td>45b</td>
<td>68</td>
<td>88</td>
</tr>
<tr>
<td>19</td>
<td>43</td>
<td>60</td>
<td>53b</td>
<td>64</td>
</tr>
<tr>
<td>20</td>
<td>28</td>
<td>56b</td>
<td>70</td>
<td>68</td>
</tr>
<tr>
<td>X</td>
<td>34.30</td>
<td>43.70</td>
<td>45.20</td>
<td>80.00</td>
</tr>
<tr>
<td>SD</td>
<td>8.25</td>
<td>19.61</td>
<td>21.12</td>
<td>8.70</td>
</tr>
</tbody>
</table>

^a Test ear thresholds for Group 1: \( \bar{x} = 1.60, \) SD = 2.17.

^b Test ear thresholds for Group 2: \( \bar{x} = 39.20, \) SD = 18.12.
<table>
<thead>
<tr>
<th>Subject</th>
<th>Test Ear</th>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1000 Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
<th>6000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>R L</td>
<td>R L</td>
<td>R L</td>
<td>R L</td>
<td>R L</td>
<td>R L</td>
</tr>
<tr>
<td>11</td>
<td>R</td>
<td>5 5</td>
<td>0 5</td>
<td>10 15</td>
<td>80 80</td>
<td>90 100</td>
<td>85 NRa</td>
</tr>
<tr>
<td>12</td>
<td>L</td>
<td>15 15</td>
<td>10 5</td>
<td>20 15</td>
<td>60 60</td>
<td>85 85</td>
<td>85 90</td>
</tr>
<tr>
<td>13</td>
<td>R</td>
<td>0 10</td>
<td>20 30</td>
<td>50 55</td>
<td>60 55</td>
<td>60 65</td>
<td>70 65</td>
</tr>
<tr>
<td>14</td>
<td>L</td>
<td>20 40</td>
<td>40 35</td>
<td>70 40</td>
<td>65 75</td>
<td>65 85</td>
<td>90 NRa</td>
</tr>
<tr>
<td>15</td>
<td>L</td>
<td>30 30</td>
<td>40 30</td>
<td>70 65</td>
<td>75 75</td>
<td>65 60</td>
<td>75 60</td>
</tr>
<tr>
<td>16</td>
<td>R</td>
<td>35 40</td>
<td>45 45</td>
<td>50 70</td>
<td>70 50</td>
<td>55 50</td>
<td>50 45</td>
</tr>
<tr>
<td>17</td>
<td>L</td>
<td>40 30</td>
<td>50 45</td>
<td>75 65</td>
<td>75 65</td>
<td>65 55</td>
<td>75 75</td>
</tr>
<tr>
<td>18</td>
<td>R</td>
<td>45 45</td>
<td>40 65</td>
<td>50 70</td>
<td>60 70</td>
<td>50 60</td>
<td>55 55</td>
</tr>
<tr>
<td>19</td>
<td>L</td>
<td>55 50</td>
<td>55 50</td>
<td>65 55</td>
<td>90 75</td>
<td>80 90</td>
<td>85 95</td>
</tr>
<tr>
<td>20</td>
<td>R</td>
<td>50 60</td>
<td>55 60</td>
<td>60 80</td>
<td>80 75</td>
<td>65 65</td>
<td>50 60</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>32.5</td>
<td>35.0</td>
<td>40.0</td>
<td>40.0</td>
<td>55.0</td>
<td>60.0</td>
</tr>
<tr>
<td>Test Ear</td>
<td></td>
<td>72.5</td>
<td>72.5</td>
<td>65.0</td>
<td>65.0</td>
<td>75.0</td>
<td>67.5</td>
</tr>
<tr>
<td>Limits</td>
<td></td>
<td>0-50</td>
<td>0-55</td>
<td>10-65</td>
<td>60-80</td>
<td>50-90</td>
<td>50-NRa</td>
</tr>
</tbody>
</table>

NR: No response at intensity limits of the audiometer.
rectifier with a 1000 Hz driving signal until the desired values of total harmonic distortion were measured on a frequency analyzer. Details of production are described by Lawson (1980) and the calibration wave forms used to obtain desired percentages of total harmonic distortion are shown in Appendix A.

**Stimulus Materials**

The stimuli consisted of orally presented passages of continuous discourse presented by a female speaker. The passages, selected from a junior high school history text, were about 10 seconds in duration and were selected on the basis of Fang's (1966, 1967) "Easy Listening Formula" (ELF). Fang concluded that an "average" ELF score below 12 is desirable for mass listenability and the six stimulus passages were rated at 9.0 or lower by two out of three independent scorers. Appendix B is a transcript of the stimulus passages.

**Signal Degradation**

The stimuli were degraded by linear rectification to produce seven levels of total harmonic distortion: 1, 10, 20, 30, 40, 50, and 60%. The level with 1% total harmonic distortion represented a "no" degradation condition.

**Test Tapes**

Two test tapes were prepared by splicing timing (leader) tape and the passages described above. Appendix C shows the presentation orders for the passages and the seven degrees of degradation of the
comparison passages. Each test tape consisted of: (a) a level cali-
bration tone, (b) spoken instructions for each task (quality and in-
telliibibility), (c) practice materials consisting of two seven-item
trials for each task, and (d) experimental materials consisting of
two seven-item trials for each task. The order of the tasks was
counterbalanced on the two tapes such that half the subjects in each
group made quality judgments first. Instructions for the listening
tasks were presented orally (via tape recording) and in writing. The
written instructions and a sample response sheet are shown in
Appendix D.

Procedures

Audiometric Screening

Subjects were evaluated to determine whether or not they met the
audiological criteria for the study and to provide reference thresh-
holds for the experimental task. Each subject was asked to sign an
informed consent release form (Appendix E), to provide case history
data, and to undergo several audiological tests. Test results and
case history information were recorded on the audiological screening
form shown in Appendix F.

Normal hearing subjects received a pure tone screening test, a
reflex decay test at 10 dB sensation level (re: acoustic reflex
threshold) at 500 and 1000 Hz and tympanometry.

Hearing impaired subjects received a threshold test (Carhart &
Jerger, 1959), an audiometric tone decay test (Olsen & Noffsinger,
1974) at 1000 and 4000 Hz and tympanometry.

The above tests were administered to each ear of all subjects and the "better" ear was designated as that ear which produced the best two-frequency pure tone average threshold. If there was no interaural difference, the right ear was arbitrarily designated as the test ear. The test ear was the same ear in which a hearing aid was worn for all but four of the hearing impaired subjects. A speech discrimination test score was obtained for the test ear by presenting the NU Auditory Test No. 6 at 40 dB sensation level (re: the two-frequency average threshold).

Visual Magnitude Estimation Training and Screening

The training and screening program was conducted individually using a rear screen slide viewer equipped with a synchronized tape player. The audio signal was presented through earphones at a comfortable loudness level.

Following a practice trial with randomly ordered squares, each subject made magnitude estimates of circle size in two subsequent trials with different randomizations of the same stimuli. To be retained for the study, a subject was required to produce a correlation (r) between estimates for the second and third trials equal to or greater than .90 and to assign the same estimate to a standard stimulus and its equivalent comparison stimulus. Lawson (1980) found that subjects who met these criteria yielded results similar to those reported by Stevens (1975).
Listening Tasks

The listening tasks, which followed the visual magnitude estimation training and screening, consisted of a practice session and an experimental session for each task (quality and intelligibility). Subjects were tested in a double-walled sound room adjacent to a single-walled control room. In the control room the output of a Sony reel-to-reel tape recorder was routed to the tape input of a two-channel speech audiometer (Grason-Stadler 162), then through the wall to a pair of TDH-49 earphones with MX 41/AR cushions located in the test room. One earphone of the pair was a "dummy." The frequency response curve for the speech audiometer with the earphones is shown in Figure 1. However, the high frequency cut-off was imposed at 4400 Hz by the computer system used to generate the test stimuli. Stimuli were presented at 40 dB above the two-frequency average threshold (Fletcher, 1950).

Subjects were instructed, orally (via tape recording) and in writing, to assign any numerical value they wished to the standard stimulus for each trial and to assign a related numerical value (a magnitude estimate) to each comparison stimulus. The instructions (shown in Appendix D) were followed by 14 practice items (two seven-item trials). In each pair, the standard passage always represented the fourth degradation level; the comparison passage represented one of the seven degradation levels. The order of comparison degradation levels varied randomly within a trial so that a given order was not repeated. The standard and comparison passages lasted approximately
Figure 1. Frequency response curve for the Grason-Stadler 162 speech audiometer with TDH-49 earphones.
10 seconds and were followed by a 5-second response interval.

Practice stimuli were followed by experimental stimuli on the same reel and the task remained the same. The entire sequence was repeated for the second task (quality or intelligibility).

Calibration of Listening Apparatus

Within-session calibration consisted of checking the signal level to the earphone prior to each experimental session. In addition, the tape recorder, speech audiometer, and earphones were checked before and after the investigation. On both occasions, the system performed within the tolerances specified by ANSI S3.6-1969.
CHAPTER III

RESULTS

Introduction

The purpose of this study was to examine speech quality magnitude estimation (SQME) functions and speech intelligibility magnitude estimation (SIME) functions of listeners with normal hearing and listeners with sensory-neural hearing loss. The magnitude estimates were obtained as a function of seven degrees of degradation by harmonic distortion.

Ten normal hearing and 10 hearing impaired subjects met all selection criteria for the study. Four additional hearing impaired subjects were eliminated. One subject was eliminated because she did not demonstrate adequate consistency in the visual magnitude estimation task, one subject was eliminated because she was not able to perform magnitude estimation with the speech stimuli, and two subjects were eliminated because they could not tolerate the presentation level for the listening task.

Data Reduction

The dependent variables considered in the study are the log geometric means of the magnitude estimates and the slopes of the functions relating the log geometric mean magnitude estimates to log stimulus values. Thus, it was necessary to reduce individual
magnitude estimates to log geometric mean magnitude estimates and slope terms. This was accomplished through the use of MAGEST, a Fortran IV computer program for analyzing magnitude estimation data (Kerst, 1978). Inputs required by the program are the perceptual magnitude estimates and the values of the stimulus magnitudes. The next two sections describe operations on the auditory stimulus values and the SQMEs. The procedure employed with the SQME data was also followed for the SIMEs.

**Stimulus Magnitudes**

Stimulus magnitudes were expressed as percent undegraded values determined by subtracting the measured percentages of total harmonic distortion from 100% total harmonic distortion. The seven levels of harmonic distortion (1, 10, 20, 30, 40, 50, and 60%) thus were expressed as the following percent undegraded values: 99, 90, 80, 70, 60, 50, and 40%. These values were entered into the MAGEST program separately for each listening task. MAGEST transforms these values by taking the natural log of each. The log stimulus values are retained in the program as independent variables to be used in subsequent processing of magnitude estimates obtained from individual subjects.

**SQMEs**

SQMEs were tabulated from individual response sheets for each subject and entered into the MAGEST program. MAGEST computes geometric mean SQMEs across trials for each subject and for each degree of
signal degradation. The geometric mean (G.M.) is defined as

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

where $X_n$ is the nth score. The geometric mean is the preferred index of central tendency (Stevens, 1975) because it is consistent with the ratio scale of measurement and because it is relatively insensitive to the effects of modulus differences across subjects or trials. MAGEST also applies the method of least squares to determine the slope and intercept of the linear equation which relates the log geometric mean SQMEs and log stimulus values.

MAGEST was executed to produce log geometric mean SQMEs and slopes for individual subjects and for groups of subjects.

Statistical Procedures

In a magnitude estimation task, the values of magnitude estimates across trials are directly affected by the numerical value of the standard stimulus, that is, the modulus. However, the slopes of the log-log functions relating magnitude estimates and stimulus values are relatively independent of the modulus chosen. Because all magnitude estimates were made in a modulus-free manner, only within-subject analyses were done on the magnitude estimates and log geometric mean magnitude estimates across trials. Both within- and between-subject analyses were done on slopes. A significance level of .05 was used in all statistical tests.
Analysis Procedures for Magnitude Estimates

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

The log geometric means were examined separately for each of the two listener groups within each listening task. This required four one-way mixed-effects analyses of variance (ANOVA) for repeated measures on the seven degradation levels (Winer, 1971). If an ANOVA showed a statistically significant trend for the log geometric means to be influenced by changes in degree of degradation, a test for linear trend (Winer, 1971) was used. The orthogonal coefficients required for this test were computed according to a procedure described by Kirk (1968) for use with unequal intervals of the independent variable. Tests for higher order trends were to be used if a linear equation failed to provide a statistically significant fit to the data, or if the linear equation failed to account for more than half the variance. The percentage (%) of variance accounted for by the linear component was estimated roughly by the following formula,

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

Description

Mean log geometric mean modulus-free magnitude estimates for Groups 1 and 2 are shown in Figure 2 for speech quality and in Figure 3 for speech intelligibility. Lines of best fit were plotted by
Figure 2. Mean log geometric mean modulus-free SQMEs for Groups 1 and 2 plotted as a function of log percent (%) undegraded by linear rectification.
Figure 3. Mean log geometric mean modulus-free SIMEs for Groups 1 and 2 plotted as a function of log percent (%) undegraded by linear rectification.
applying the method of least squares. The group functions for each listening task appear to be relatively linear in shape and the magnitudes of the log geometric mean magnitude estimates increase with decreasing stimulus degradation.

Table 3 summarizes the mean slopes, standard deviations, and ranges for listener groups as a function of listening tasks. As before, slopes were obtained by applying the method of least squares to the log geometric mean magnitude estimates and log stimulus values. The smallest slope was obtained for Group 2 on quality judgments; the largest slope was obtained for Group 1 on quality judgments. Figure 4 shows the mean slopes for groups plotted as a function of listening tasks. Group 2 produced the largest slope for each listening task, quality and intelligibility. In Figure 5 the mean slopes for listening tasks are plotted as a function of groups. The slope for quality is larger than the slope for intelligibility for Group 1; the slope for intelligibility is larger than the slope for quality for Group 2.

Table 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Quality</th>
<th></th>
<th>Intelligibility</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>SD</td>
<td>Range</td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td>1.010</td>
<td>.565</td>
<td>2.039</td>
<td>.875</td>
</tr>
<tr>
<td>2</td>
<td>.596</td>
<td>.647</td>
<td>2.435</td>
<td>.622</td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td>.803</td>
<td>.625</td>
<td>2.763</td>
<td>.748</td>
</tr>
</tbody>
</table>
Figure 4. Mean slopes for groups plotted as a function of listening tasks.
Figure 5. Mean slopes for listening tasks plotted as a function of groups.
Analysis

Presence of Trends

Tables 4 through 7 summarize the results of analyses of variance in log geometric mean magnitude estimates across trials as a function of log degradation levels for each listening task. Group 1 data are shown for quality judgments (Table 4) and intelligibility judgments (Table 5); Group 2 data are also shown for quality judgments (Table 6) and intelligibility judgments (Table 7). In each case, the observed F ratio was significant, suggesting a statistically significant trend for the log geometric mean magnitude estimates to be influenced by changes in degree of degradation.

Table 4

Summary of Analysis of Variance in Log Geometric Mean SQMEs Across Trials for Group 1 as a Function of Log Percent Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>F for α = .05</th>
<th>α for Fobs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>6.8612</td>
<td>9</td>
<td>.7624</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td>2.0719</td>
<td>60</td>
<td>.0345</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degradation levels</td>
<td>1.4788</td>
<td>6</td>
<td>.2465</td>
<td>22.439</td>
<td>2.04*</td>
<td>.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>.5931</td>
<td>54</td>
<td>.0110</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8.9330</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*From F distribution table (Winer, 1971).
Table 5

Summary of Analysis of Variance in Log Geometric Mean SIMEs Across Trials for Group 1 as a Function of Log Percent Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>α for Fobs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>6.916</td>
<td>9</td>
<td>.7684</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td>1.6075</td>
<td>60</td>
<td>.0268</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degradation levels</td>
<td>.9475</td>
<td>6</td>
<td>.1579</td>
<td>12.422</td>
<td>2.04* .0001</td>
</tr>
<tr>
<td>Residual</td>
<td>.6599</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8.5234</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*From F distribution table (Winer, 1971).

Table 6

Summary of Analysis of Variance in Log Geometric Mean SQMEs Across Trials for Group 2 as a Function of Log Percent Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>α for Fobs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>8.6835</td>
<td>9</td>
<td>.9648</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td>1.7630</td>
<td>60</td>
<td>.0294</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degradation levels</td>
<td>.4722</td>
<td>6</td>
<td>.0787</td>
<td>3.292</td>
<td>2.04* .008</td>
</tr>
<tr>
<td>Residual</td>
<td>1.2909</td>
<td>54</td>
<td>.0239</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10.4465</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*From F distribution table (Winer, 1971).
Table 7

Summary of Analysis of Variance in Log Geometric Mean SIMEs Across Trials for Group 2 as a Function of Log Percent Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>F for α = .05</th>
<th>α for Fobs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>5.7440</td>
<td>9</td>
<td>.6382</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td>1.3708</td>
<td>60</td>
<td>.0228</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degradation levels</td>
<td>.5047</td>
<td>6</td>
<td>.841</td>
<td>5.244</td>
<td>2.04*</td>
<td>.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>.8661</td>
<td>54</td>
<td>.0160</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>7.1147</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*From F distribution table (Winer, 1971).

Nature of Trends

Tables 8 through 11 summarize the results of tests to determine whether a significant portion of the trends detected by the analysis of variance could be accounted for by a linear equation. Group 1 data are shown for speech quality magnitude estimates (Table 8) and for speech intelligibility magnitude estimates (Table 9). Group 2 data are shown for speech quality magnitude estimates (Table 10) and for speech intelligibility magnitude estimates (Table 11). In each case, the observed F ratio for linear trend was significant, suggesting that a linear equation provides a statistically significant fit to the data. Also, the linear component of the trend in each case accounted for over 50% of the variance in log geometric mean.
Table 8

Results of Test for Linear Trend in Log Geometric Mean SQMEs for Group 1 as a Function of Log Percent (%) Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>F for α = .05</th>
<th>α for Fobs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear trend</td>
<td>1.22</td>
<td>1</td>
<td>1.22</td>
<td>214.04</td>
<td>4.003*</td>
<td>.0001</td>
</tr>
<tr>
<td>Deviation from linear trend</td>
<td>.33</td>
<td>59</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Interpolated from F distribution table (Winer, 1971).

Table 9

Results of Test for Linear Trend in Log Geometric Mean SIMEs for Group 1 as a Function of Log Percent (%) Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>F for α = .05</th>
<th>α for Fobs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear trend</td>
<td>.92</td>
<td>1</td>
<td>.92</td>
<td>76.67</td>
<td>4.003*</td>
<td>.0001</td>
</tr>
<tr>
<td>Deviation from linear trend</td>
<td>.69</td>
<td>59</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Interpolated from F distribution table (Winer, 1971).
Table 10

Results of Test for Linear Trend in Log Geometric Mean SQMEs for Group 2 as a Function of Log Percent (%) Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>F for α = .05</th>
<th>α for Fobs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear trend</td>
<td>.38</td>
<td>1</td>
<td>.38</td>
<td>16.52</td>
<td>4.003*</td>
<td>.0001</td>
</tr>
<tr>
<td>Deviation from linear trend</td>
<td>1.38</td>
<td>59</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Interpolated from F distribution table (Winer, 1971).

Table 11

Results of Test for Linear Trend in Log Geometric Mean SIMEs for Group 2 as a Function of Log Percent (%) Undegraded by Linear Rectification

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>F for α = .05</th>
<th>α for Fobs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear trend</td>
<td>.30</td>
<td>1</td>
<td>.30</td>
<td>16.72</td>
<td>4.003*</td>
<td>.0001</td>
</tr>
<tr>
<td>Deviation from linear trend</td>
<td>1.07</td>
<td>59</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Interpolated from F distribution table (Winer, 1971).
magnitude estimates due to log degradation levels. Table 12 summarizes the approximate percentages of variance that can be attributed to the linear component of each trend. Because of these results, tests for higher order trends were not employed.

Table 12

Approximate Percentages of Variance in Log Geometric Mean Magnitude Estimates Due to Log Degradation Levels That Could Be Accounted for by a Linear Equation

<table>
<thead>
<tr>
<th>Group</th>
<th>Listening Task</th>
<th>% Variance³</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Normal)</td>
<td>Quality</td>
<td>82.43</td>
</tr>
<tr>
<td>1</td>
<td>Intelligibility</td>
<td>97.10</td>
</tr>
<tr>
<td>2 (Impaired)</td>
<td>Quality</td>
<td>80.47</td>
</tr>
<tr>
<td>2</td>
<td>Intelligibility</td>
<td>59.64</td>
</tr>
</tbody>
</table>

Note. Percentages (%) are shown as a function of group and listening task.

³% variance = 100 \( \frac{SS_{1in}}{SS_{degradation \ level}} \)

Analysis Procedures for Slopes

Description

The slopes for the log-log functions were examined for statistically significant differences as a function of groups, listening tasks, and interactions of groups and listening tasks. A two-way ANOVA (2x2) for repeated measures (Winer, 1971) was performed and the
exact probabilities of Type 1 errors were estimated for all F ratios.

**Analysis**

The results of the two-way analysis of variance as a function of the two listener groups and two listening tasks are summarized in Table 13. Significant F ratios are those which exceed the critical value of F at an alpha level of .05. The main effects of groups, listening tasks, and groups-by-listening tasks were not significant. The main effect for groups produced the largest F value while the main effect for listening tasks produced the smallest F value.

**Reliability**

Pearson product-moment correlation coefficients (Hays, 1963) were used to assess within-session reliability of the visual and auditory magnitude estimation tasks. The strength of each correlation was examined for statistical significance (Byrkit, 1972) and described according to the guidelines suggested by Silverman (1977).

**Visual Magnitude Estimation**

Pearson product-moment correlation coefficients (r) between each subject's Trial 2 and Trial 3 visual magnitude estimates of circle size (in²) are displayed in Table 14. All coefficients were significant beyond the .05 level (df = 5; r_critical = .754). Coefficients ranged from .973 to 1.000 for Group 1 and from .910 to .999 for Group 2. These results suggest a very high degree of within-subject reliability for both groups.
Table 13

Results of a Two-Way Analysis of Variance in Slopes as a Function of Two Listener Groups (Group 1 and Group 2) and Two Listening Tasks (Quality and Intelligibility)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>F for (\alpha = .05)</th>
<th>(\alpha) for (F_{obs})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>13.2858</td>
<td>19</td>
<td>.6993</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groups</td>
<td>1.1112</td>
<td>1</td>
<td>1.1112</td>
<td>1.643</td>
<td>4.38*</td>
<td>.216</td>
</tr>
<tr>
<td>Subjects within groups</td>
<td>12.1746</td>
<td>18</td>
<td>.6764</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within subjects</td>
<td>1.9279</td>
<td>20</td>
<td>.0964</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening tasks</td>
<td>.0301</td>
<td>1</td>
<td>.0301</td>
<td>.295</td>
<td>4.35*</td>
<td>.593</td>
</tr>
<tr>
<td>Groups X listening tasks</td>
<td>.0649</td>
<td>1</td>
<td>.0649</td>
<td>.637</td>
<td>4.35*</td>
<td>.435</td>
</tr>
<tr>
<td>Listening tasks X subjects</td>
<td>1.8324</td>
<td>18</td>
<td>.1018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>15.2137</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*From F distribution table (Winer, 1971).
Table 14

Pearson Product-Moment Correlation Coefficients (r) Between the Visual Magnitude Estimates of Circle Size (in²) for Each Subject's Trial 2 and Trial 3 Stimuli

<table>
<thead>
<tr>
<th>Group 1 (Normal)</th>
<th>Group 2 (Impaired)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>r</td>
</tr>
<tr>
<td>1</td>
<td>.997</td>
</tr>
<tr>
<td>2</td>
<td>.996</td>
</tr>
<tr>
<td>3</td>
<td>.991</td>
</tr>
<tr>
<td>4</td>
<td>.980</td>
</tr>
<tr>
<td>5</td>
<td>.982</td>
</tr>
<tr>
<td>6</td>
<td>.997</td>
</tr>
<tr>
<td>7</td>
<td>.999</td>
</tr>
<tr>
<td>8</td>
<td>.973</td>
</tr>
<tr>
<td>9</td>
<td>.997</td>
</tr>
<tr>
<td>10</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Auditory Magnitude Estimation

Table 15 lists individual subject and group "average" correlation coefficients between magnitude estimates for the first and second trials (n = 7) for each listening task. Group "average" correlations also are shown graphically in Figure 6. Significant coefficients at the .05 level were those which equaled or exceeded .754. For Group 1, significant correlations were found for the quality judgments of three subjects and for the intelligibility judgments of four subjects. The "average" coefficients for Group 1 were .63 for the quality task and .71 for the intelligibility task. These averages are not statistically significant but suggest moderate within-subject reliability of magnitude estimates for Group 1 under each listening task. For Group 2, three of the individual correlations were significant for quality judgments and none were significant for intelligibility judgments. "Average" coefficients for Group 2 were .53 for the quality task and .45 for the intelligibility task, suggesting weak within-subject reliability of magnitude estimates for each listening task.

Figure 7 shows the within-group correlation coefficients between Trial 1 and Trial 2 slopes for each listener group (n = 10) for each listening task. Significant coefficients were those which equaled or exceeded .632. For Group 1 quality judgments the correlation between Trial 1 and Trial 2 slopes was .63 and for Group 1 intelligibility judgments, the correlation between slopes was .64. These correlations suggest moderate within-group reliability for each listening task.
Table 15

Within-Subject Correlation Coefficients (Pearson r) Between Seven Trial 1 and Trial 2 Magnitude Estimates From Each Task for Each of 20 Subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>Listening Task</th>
<th></th>
<th>Subject</th>
<th>Listening Task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SQME</td>
<td>SIME</td>
<td></td>
<td>SQME</td>
</tr>
<tr>
<td>Group 1 (Normal)</td>
<td></td>
<td></td>
<td>Group 2 (Impaired)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.17</td>
<td>.83</td>
<td>11</td>
<td>.80</td>
</tr>
<tr>
<td>2</td>
<td>.88</td>
<td>.90</td>
<td>12</td>
<td>-.27</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.49</td>
<td>13</td>
<td>.23</td>
</tr>
<tr>
<td>4</td>
<td>.56</td>
<td>.73</td>
<td>14</td>
<td>-.18</td>
</tr>
<tr>
<td>5</td>
<td>.88</td>
<td>.94</td>
<td>15</td>
<td>.00</td>
</tr>
<tr>
<td>6</td>
<td>.50</td>
<td>.07</td>
<td>16</td>
<td>.77</td>
</tr>
<tr>
<td>7</td>
<td>.48</td>
<td>.69</td>
<td>17</td>
<td>.70</td>
</tr>
<tr>
<td>8</td>
<td>.68</td>
<td>.93</td>
<td>18</td>
<td>.43</td>
</tr>
<tr>
<td>9</td>
<td>.68</td>
<td>-.22</td>
<td>19</td>
<td>.64</td>
</tr>
<tr>
<td>10</td>
<td>.77</td>
<td>.41</td>
<td>20</td>
<td>.79</td>
</tr>
</tbody>
</table>

$\bar{x}_r^a = .63 \quad .71 \quad \bar{x}_r^a = .53 \quad .45$

Note. "Average" coefficients were determined for groups within tasks.

$a$"Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963).
Figure 6. "Average" within-subject test-retest correlation coefficients between seven Trial 1 and Trial 2 SQMEs and SIMEs for two groups of 10 subjects. "Average" correlations are Fisher's Z to r transformations interpolated from mean Fisher's r to Z transformations (Hays, 1963). The dashed horizontal line denotes the significance criterion (r = .754).
Figure 7. Within-group correlation coefficients between Trial 1 and Trial 2 slopes for each group under each condition. The dashed horizontal line denotes the significance criterion (r = .632).
task. For Group 2, the correlation between Trial 1 and Trial 2 slopes was .44 for the quality task, suggesting weak reliability, and .72 for the intelligibility task, suggesting moderate reliability.
CHAPTER IV

DISCUSSION

Introduction

Chapter III described the results produced by 10 subjects with normal hearing (Group 1) and 10 subjects with sensory-neural hearing impairment (Group 2) in a speech quality magnitude estimation task and in a speech intelligibility magnitude estimation task. In both tasks, modulus-free magnitude estimates were obtained as a function of seven degrees of signal degradation by harmonic distortion.

Two dependent variables were derived from the individual magnitude estimates for each task: (a) log geometric mean magnitude estimates and (b) slopes for the lines relating log geometric mean magnitude estimates to log stimulus values. The log geometric mean magnitude estimates were examined for statistically significant differences as a function of degrees of degradation for each listening group and task. When differences were found, the lowest order equation required to provide a statistically significant fit to the data was determined. The slopes were analyzed to find statistically significant differences as a function of listener group, listening task, or interaction of group and task. The findings and their implications are discussed below.
Log Geometric Mean Magnitude Estimates

The first experimental question asked if there was a statistically significant trend for the log geometric mean magnitude estimates for each task (quality and intelligibility) to be influenced by change in degree of degradation for each listener group. The results obtained indicate that such a trend does exist. Statistically significant differences were found in the log geometric means for Group 1 and Group 2 as a function of degree of degradation for each of the listening tasks.

The second experimental question concerned determination of the lowest order equation required to provide a statistically significant fit to the data for each listening group under each listening task. A first order equation provides a statistically significant fit to the data from each task for each listening group and accounts for a large percentage of the variance in log geometric means as a function of log degradation levels. It can be concluded that a positive linear relation exists between log perceptual values and log stimulus values for Group 1 and Group 2 under each listening task.

The linear relation of the log geometric mean magnitude estimates to log percent undegraded by harmonic distortion suggests that magnitude estimates are related to changes in the electroacoustic characteristics of the signal by a power function. These results are in agreement with those of Lawson (1980) and Barry and Kidd (1981) who also found a linear relation between log perceptual judgments and log harmonic distortion.
Slopes

The third experimental question asked whether the slopes of the log-log functions showed statistically significant differences as a function of listener group, listening task, or interaction of group and task. The results of the study revealed no statistically significant differences among the slopes for any of these functions. The means, standard deviations, and ranges of the slopes are shown in Table 3 (page 31) as a function of groups and tasks. For Group 1, the mean slope for quality was larger than the mean slope for intelligibility, although the difference was not statistically significant. For Group 2, the mean slope for intelligibility was larger than the mean slope for quality, suggesting that the effects of signal distortion may differ for the hearing impaired. Again, the difference was not statistically significant.

The Group 1 mean slope for quality judgments, 1.010, indicates that growth of sensation and growth of degradation were equal for that experimental condition. Slopes for the remaining three experimental conditions were less than 1; therefore, growth of sensation was slower than growth of degradation for these conditions.

For both listening tasks, the mean slope for Group 1 was larger than the mean slope for Group 2, suggesting that Group 1 was more sensitive to changes in degradation level than was Group 2. Again, the differences in the slopes were not significant.

These results differ from those of Lawson (1980) in a similar study involving magnitude estimates of quality for speech degraded by
harmonic distortion. He found a statistically significant difference between slopes produced by listeners with normal hearing and those produced by listeners with sensory-neural hearing loss. However, both his study and the present one found the slopes produced by normal hearing listeners to be greater than those produced by hearing impaired listeners. The mean slope for hearing impaired listeners was .73 in the Lawson (1980) study and .60 in the present one. However, on the average, subjects in this study exhibited greater hearing losses than those in the Lawson (1980) study. Median threshold differences ranged from 5 dB at 4000 and 6000 Hz to 20 dB at 2000 Hz. The value of the mean slope for normal hearing listeners' quality judgments was slightly greater than one in each study. These findings differ from those of Barry and Kidd (1981) who found a mean slope of .75 for magnitude estimates by normal hearing listeners. However, Barry and Kidd (1981) used a wider range of harmonic distortion values (10-100%) than did the Lawson (1980) study and the present study (1-60%). Stevens (1975) noted that reducing the range of the stimulus magnitude tends to reduce the slope of the function.

Reliability

Visual Magnitude Estimation

The reliability of each subject's Trial 2 and Trial 3 visual magnitude estimates of circle size was assessed by Pearson product-moment correlation coefficients. All subjects produced a statistically significant strong positive correlation ($r > .90$) during the
training task, suggesting that each subject was able to reliably perform a magnitude estimation task. These findings indicate that the results of the auditory magnitude estimation task should not have been biased by differences in the subject's magnitude estimation scaling abilities.

**Auditory Magnitude Estimation**

Pearson product-moment correlation coefficients (r) were used to assess the within-subject reliability within experimental listening sessions. Within-session correlation coefficients were obtained between each subject’s Trial 1 and Trial 2 magnitude estimates for each task (quality and intelligibility). For both listening tasks, "average" individual correlations suggested moderate reliability for the normal hearing subjects and weak reliability for the hearing impaired subjects. No "average" correlations met the significance criterion.

For each group, within-group correlation coefficients were determined between Trial 1 and Trial 2 slopes for each listening task. For Group 1, reliability was found to be moderate for each task. For Group 2, reliability was weak for the quality task and moderate for the intelligibility task. However, all correlations for slopes, except the Group 2 correlation for quality, met significance criterion.

The within-session reliability of the individual magnitude estimates and slopes for trials is important because the dependent variables (i.e., the log geometric mean magnitude estimates and the mean slopes for groups) were derived from the magnitude estimates. Therefore, the weak to moderate reliability of the magnitude estimates is
reflected in the weak to moderate reliability of the within-group Trial 1 and Trial 2 slopes and may underlie the failure to find statistically significant differences between the mean slopes as a function of group, task, or interaction of group and task.

The variability of the magnitude estimates produced by each listener group may have contributed to the variability of the slopes and large variability of slopes within the experimental conditions may have precluded finding statistically significant differences between the conditions. Nevertheless, the present study presents some interesting implications for future research.

Implications

Reliability

The most obvious question which arises from the current study is the cause of the low reliability factor. Recall that all subjects underwent a visual magnitude estimation training procedure and produced Pearson product-moment correlation coefficients (r) greater than .90 for magnitude estimates of circle size. The training procedure was developed by Lawson (1980) who found that it adequately prepared subjects to perform magnitude estimations for auditory stimuli. His results are in agreement with those of Stevens (1975) who found that a simple task such as judgment of line length or circle size was helpful in initiating untrained subjects. In the present study, reliable performance of visual magnitude estimation did not appear to transfer to the auditory task, suggesting the need for further...
research in this area.

Perceptual Attributes

Quality and intelligibility were clearly defined as the perceptual attributes to be judged. However, no significant differences were found between the quality judgments and the intelligibility judgments. Intuitively, quality appears to differ from intelligibility but, statistically, that difference remains to be demonstrated. Disregarding the reliability problem, it may be that intelligibility constitutes such a major part of quality that a more strictly controlled experiment would be necessary to demonstrate the difference. It is also possible that magnitude estimation is not a sensitive enough measure to distinguish between the two.

Sex of the Talker

The speech stimuli were recorded by a female because it has been reported that female speakers seem to accentuate quality differences (Punch & Ciechanowski, 1977). While differences in both quality and intelligibility as a function of harmonic distortion were significant, differences between quality and intelligibility were not. The question arises whether or not such differences would be demonstrated with a male talker. Punch and Ciechanowski (1977) found that only male voice stimuli resulted in acceptably reliable quality judgments with dysacusis listeners. Although Lawson (1980) obtained reliable quality judgments with a female talker, the present study did not. Since the present study presented selected stimuli from the Lawson
study through a different electroacoustic system, it is possible that
differences in results could be partially accounted for by the inter-
action of voice and electroacoustic system. Witter and Goldstein
(1971) implied such an interaction, reporting that preferences for
hearing aids differed both as a function of the talker's sex and as a
function of the particular pair of aids being compared. In addition
to these issues, variables related to spectral characteristics of
male versus female speech should be examined.

**Frequency Bandwidth**

The frequency bandwidth of the auditory stimuli was a function
of the apparatus used to generate and present it. It was noted that
the upper frequency cut-off was imposed at 4400 Hz by the computer
used to generate the speech stimuli. According to Fletcher (1953),
the system would transmit neither the main frequency bands of four
consonants (f, s, voiced th, and unvoiced th) nor the second fre-
quency bands of four consonants (sh, ch, z, and v). Those sounds
whose main frequency bands would not be transmitted comprise from
9 to 13% of the sounds in the passages with a mean of 10.33%. How-
ever, such an analysis does not consider contextual and linguistic
factors in speech perception. Probability theory provides a more
realistic method for examining the efficiency of the transmission
system. French and Steinberg (1947) divided the acoustic spectrum
into 20 frequency bandwidths of equal importance to recognition of
speech sounds. The probability of phoneme recognition (P_{x/20}) asso-
ciated with a system which transmits a combination of the frequency
bands \( (x) \) is given by,

\[
P_{x/20} = 1 - (1 - P_f)^{x/20}
\]

where \( P_f \) is the given value of phoneme recognition probability (98.5%) when 100% of the acoustic information is present. According to this formula, the probability of phoneme recognition in the current study was 97.8%. Future studies should examine the relationship between phoneme recognition probability and perceptual judgments of speech.

**Laterality**

The left hemisphere in humans has long been considered dominant for language functions. The dominance prevails in the majority of left-handed people as well as right-handed ones (Eisenson, 1973; Herron, 1980). Penfield and Roberts (1966) reported that less than 10% of left-handed and less than 1% of right-handed people have some representation of speech in the right hemisphere. Thus, regardless of handedness, the vast majority of people have dominance localized in the left cerebral hemisphere.

According to Kimura (1961), the crossed auditory pathways in humans are stronger, or more numerous, than the uncrossed ones, and the ear opposite the dominant hemisphere is more efficient in processing spoken material. Therefore, the majority of people perform speech-related tasks better when the signal is presented to the right ear than to the left ear. The current experiment was monotic and selection of the test ear was a function of hearing acuity. Two of the normal hearing subjects and five of the hearing impaired subjects...
were tested in the left ear, implying that the speech may have been received primarily by the right or nondominant (for language) hemisphere. The significance of this should be interpreted with caution because, as Eisenson (1973) pointed out, positive statements about localization are still assumptions and inferences. Nevertheless, the interaction of perceptual task, hearing loss, test ear, and hemispheric dominance represents an interesting area for future study.
CHAPTER V

SUMMARY AND CONCLUSIONS

Purpose

This study was designed to examine psychophysical functions obtained on listeners with normal hearing and listeners with sensory-neural hearing impairment. Magnitude estimates for speech quality and speech intelligibility were obtained on the two groups of listeners as a function of seven degrees of harmonic distortion. The log geometric means of two within-session magnitude estimates were plotted as a function of log degree of degradation for the two listener groups under each listening task. The log-log functions were analyzed:

(a) to determine whether log geometric mean magnitude estimates differed as a function of changes in log degree of degradation for each listening group under each listening task;

(b) to identify the lowest order equation required to provide a satisfactory fit to the log-log functions; and

(c) to determine whether differences existed among the slopes of the log-log functions as a function of listener group, listening task, or interaction of group and task.
Experimental Design

Subjects

Subjects for the study were 20 adult volunteers who were divided into two groups. Group 1 consisted of 10 normal hearing listeners with a mean age of 31.10 years, a mean two-frequency average hearing threshold of 1.60 dB in the test ear, and a mean speech discrimination score of 99.20% in the test ear. Group 2 consisted of 10 sensorineurally impaired hearing listeners with a mean age of 34.30 years, a mean two-frequency average hearing threshold of 39.20 dB in the test ear, and a mean speech discrimination score of 80.00% in the test ear. All of the subjects in Group 2 were hearing aid users with length of use ranging from 3 months to 23 years.

Stimuli

Visual Training and Screening Stimuli

The visual stimuli consisted of three sets of seven pairs of geometric forms presented side by side on 2" by 2" slides. The first set of seven pairs consisted of squares, while the last two sets of seven pairs consisted of circles.

Auditory Stimuli

The auditory stimuli were tape recordings of six 10-second connected speech samples recorded in pairs. Each pair consisted of a standard stimulus followed by one of seven comparison stimuli. The
stimuli were degraded by linear rectification to produce seven levels of total harmonic distortion: 1, 10, 20, 30, 40, 50, and 60%. The standard stimulus always represented the middle (fourth) degradation level.

**Procedures**

All subjects participated in a hearing screening, visual magnitude estimation training and screening, and the experimental listening tasks.

**Hearing Screening**

The hearing screening consisted of a brief history, pure tone air conduction testing, tympanometry, reflex decay or tone decay testing, and word discrimination testing in the test ear. Pure tone air conduction testing involved a screening test for the normal group and a threshold test for the impaired group. A reflex decay test was administered to the normal subjects and an audiometric tone decay test was administered to the impaired subjects.

**Visual Magnitude Estimation Training and Screening**

Each subject was presented with three sets of randomly ordered pairs of geometric forms, which were accompanied by tape recorded instructions. The subject assigned a numerical value to the magnitude of the standard stimulus in each pair and then estimated the magnitude of the comparison stimulus relative to that of the standard.
The first set, which consisted of squares, was a practice trial; the second and third sets, which consisted of circles, provided a test-retest reliability check.

**Listening Tasks**

Each subject participated in a practice session and an experimental session for each listening task. Practice stimuli and experimental stimuli consisted of 14 pairs of passages (two seven-item trials). The order of degradation levels varied randomly within each trial so that a given order was not repeated. The order of the tasks was counterbalanced such that half the subjects in each group made quality judgments first and half made intelligibility judgments first. Listeners assigned magnitude estimates to the auditory standard stimulus and comparison stimuli in the same manner used to make visual magnitude estimates in the visual task.

**Dependent Variables**

Two dependent variables were derived from the auditory magnitude estimates: (a) log geometric means of the magnitude estimates and (b) the slopes of the functions relating the log geometric mean magnitude estimates to log stimulus values.

**Findings**

Findings of the study provided the following answers to the experimental questions:
1. There was a statistically significant trend for the log geometric mean magnitude estimates for each task (quality and intelligibility) to be influenced by change in degree of degradation for each listener group.

2. The lowest order equation required to provide a satisfactory (i.e., statistically significant) fit to the data for each listening group under each listening task was a first order, or linear, equation.

3. There was no statistically significant difference among the slopes of the log-log functions as a function of: (a) listener group, (b) listening task, or (c) interaction of group and task.

Conclusions

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

1. Performance on the visual training and screening program may not be an adequate predictor of performance on auditory magnitude estimation tasks.

2. Given the relatively large variance of magnitude estimates and individual slopes, the failure to find differences between the slopes of the log-log functions does not preclude the existence of such differences.

3. Additional research is needed on quality and intelligibility judgments of speech by normal hearing and hearing impaired individuals.
APPENDICES
Appendix A

Calibration Waveforms Used to Obtain Desired Percentages of Total Harmonic Distortion

Scale:

2v/unit

0.5 msec/unit
Appendix B

Transcript of Six Stimulus Passages

Practice Standard B

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

Practice Comparison B

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

Experimental Standard A

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

Experimental Comparison A

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

Experimental Standard C

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

Experimental Comparison C

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

Appendix C

Crossbreak Matrix: Presentation Orders for Passages and Random Orders for Comparison Degradation Levels

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Trials</th>
<th>Pairings of Six Stimulus Passages&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Order of Seven Comparison Degradation Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tape 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1- 5</td>
<td>(Quality Instructions)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>PSB &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2135764</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>PSB &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2653741</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ESA &amp; ECA&lt;sub&gt;1&lt;/sub&gt;-ECA&lt;sub&gt;7&lt;/sub&gt;</td>
<td>7613542</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>ESC &amp; ECC&lt;sub&gt;1&lt;/sub&gt;-ECC&lt;sub&gt;7&lt;/sub&gt;</td>
<td>1743652</td>
</tr>
<tr>
<td></td>
<td>(Intelligibility Instructions)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>PSB &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2574316</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>PSB &amp; PCB&lt;sub&gt;7&lt;/sub&gt; &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>4316572</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ESC &amp; ECC&lt;sub&gt;1&lt;/sub&gt;-ECC&lt;sub&gt;7&lt;/sub&gt;</td>
<td>4267153</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>ESA &amp; ECA&lt;sub&gt;1&lt;/sub&gt;-ECA&lt;sub&gt;7&lt;/sub&gt;</td>
<td>7631524</td>
</tr>
<tr>
<td><strong>Tape 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-10</td>
<td>(Intelligibility Instructions)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>PSB &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2135764</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>PSB &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2653741</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ESA &amp; ECA&lt;sub&gt;1&lt;/sub&gt;-ECA&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2143657</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>ESC &amp; ECC&lt;sub&gt;1&lt;/sub&gt;-ECC&lt;sub&gt;7&lt;/sub&gt;</td>
<td>4723156</td>
</tr>
<tr>
<td></td>
<td>(Quality Instructions)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>PSB &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2574316</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>PSB &amp; PCB&lt;sub&gt;1&lt;/sub&gt;-PCB&lt;sub&gt;7&lt;/sub&gt;</td>
<td>4316572</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ESC &amp; ECC&lt;sub&gt;1&lt;/sub&gt;-ECC&lt;sub&gt;7&lt;/sub&gt;</td>
<td>6512374</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>ESA &amp; ECA&lt;sub&gt;1&lt;/sub&gt;-ECA&lt;sub&gt;7&lt;/sub&gt;</td>
<td>3152647</td>
</tr>
</tbody>
</table>

<sup>a</sup>PSB: practice standard passage B
PCB<sub>1</sub>-PCB<sub>7</sub>: 7 degradation levels of practice comparison passage B
ESA: experimental standard passage A
ECA<sub>1</sub>-ECA<sub>7</sub>: 7 degradation levels of experimental comparison passage A
ESC: experimental standard passage C
ECC<sub>1</sub>-ECC<sub>7</sub>: 7 degradation levels of experimental comparison passage C

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Appendix D

Written Instructions and Sample Response Sheet

Instructions for First Listening Task

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

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

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

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

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

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

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

Look at your response sheets. After each item you will have about 5 seconds to write a number in the appropriate box. Start with the blocks for Trial 1 in the left-hand column and work from top to bottom as you did before. You should have two response sheets and each sheet should have two columns.

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

Instructions for Second Listening Task

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

You will again be hearing pairs of passages. The first passage in each pair will always be the standard passage and the second passage will always be the comparison passage. In this task, the intelligibility/quality of the comparison passage may change from time to time.

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

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

After hearing a trial of seven pairs of passages, you will hear a new trial beginning with a new standard passage. Just follow the same procedure used before.
We're almost ready to start. Please pay careful attention to what you hear. Remember, your job is to recall the number you gave the standard and to assign a number to the comparison that represents the intelligibility/quality of the comparison relative to the intelligibility/quality of the standard.

Look at your response sheet. After each item you will have about 5 seconds to write a number in the appropriate box. Start with the blocks for Trial 1 in the left-hand column and work from top to bottom as you did before. You should have two response sheets and each sheet should have two columns.

If you want to reread these instructions or ask a question, inform the experimenter. Also, inform the experimenter when you are ready to begin the task.
<table>
<thead>
<tr>
<th>Subject No.</th>
<th>Date</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Stimulus Sets**

<table>
<thead>
<tr>
<th>Trial 1</th>
<th>Trial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>
Appendix E

Informed Consent Release Form

1. I, ___________________________, freely and voluntarily consent to serve as a subject in a scientific study of speech perception conducted by Dr. Gary Lawson, Ms. Theresa Smith, and other student assistants.

2. I understand that the purpose of the study is to determine the relations between the perceived quality and perceived intelligibility of speech which may be of future clinical usefulness.

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

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

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

6. I have had the opportunity to ask questions about the nature and purpose of the study, and I understand that upon completion of the study, and at my request, I can obtain additional explanation about the study.

Date: ____________________ Signed: __________________________
Appendix F

AUDIOLOGICAL SCREENING FORM

Project: Date: Time: Examiner:

Subject Identification
Name: Subject No.
Birthdate: Age:

History
Recent onset of hearing loss? Yes No
Vertigo? Yes No
Otoxic surgery? Yes No
Handedness? L R

Impaired hearing if only
Family history of genetic hearing loss? Yes No
Hearing aid use? L R

Length of hearing aid use?

Test Results

Pure Tone Air Conduction Thresholds (dB HL)
Frequency (Hz) 250 500 1000 2000 4000 6000 2-cone Ave.

Ear  R ___ ___ ___ ___ ___ ___ ___ L ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ___ ...
BIBLIOGRAPHY


