1981

Spondee thresholds as a function of psychophysical method and increment size| A cost-effectiveness study

Lynn V. Harris

The University of Montana

SPONDEE THRESHOLD AS A FUNCTION OF PSYCHOPHYSICAL

METHOD AND INCREMENT SIZE: A COST-EFFECTIVENESS STUDY

by

Lynn V. Harris

B.A., Communication Sciences and Disorders
University of Montana, 1979

An Abstract

Of a thesis submitted in partial fulfillment of requirements for the degree of Master of Arts in the Department of Communication Sciences and Disorders in the Graduate School of the University of Montana

October 1981

Approved by: ______________________
Chairman, Board of Examiners

____
Dean, Graduate School

10/28/81
Date

ABSTRACT

Harris, Lynn V., M.A., December 1981 Communication Sciences and Disorders

Spondee thresholds: a cost-effectiveness study (133 pp.)

Director: Michael J.M. Raffin, Ph.D.

Thesis approved: Michael J M Raffin ________________

Spondee thresholds using ascending, descending and bracketing methods for each of two increment sizes (2 dB and 5 dB) were determined for each of 60 normal hearing subjects. Each psychophysical method and increment size combination was administered twice resulting in a total of 12 spondee thresholds per subject. These 12 measurements were compared to the two- and three-frequency pure-tone average, and the number of spondees needed for each trial also was computed. Results indicated that bracketing and descending methods for either 2 or 5 dB steps were approximately equal. The descending 2 dB method required fewer spondees than any other method. Correlations between spondee thresholds and the pure-tone averages did not exceed 0.6292. Discrepancies between the spondee threshold and the pure-tone averages contradict reports in the literature.

SPONDEE THRESHOLDS AS A FUNCTION OF PSYCHOPHYSICAL METHOD AND INCREMENT SIZE: A COST-EFFECTIVENESS STUDY

by

Lynn V. Harris

B.S. Communication Sciences and Disorders
University of Montana, 1979

A Thesis

Submitted in partial fulfillment of the
requirements for the degree of Master of Arts in the
Department of Communication Sciences and Disorders
in the Graduate School of
the University of Montana

October 1981

Thesis Director: Michael J.M. Raffin, Ph.D.

Graduate School
The University of Montana
Missoula, Montana

CERTIFICATE OF APPROVAL

M.A. THESIS

This is to certify that the M.A. Thesis of

Lynn V. Harris

has been approved by the Board of Examiners
for the thesis requirements for the degree of
Master of Arts in the Department of
Communication Sciences and Disorders

Dean, Graduate School: ______________________

Board of Examiners: ______________________
Michael J.M. Raffin, Thesis Director
____
Barbara A. Bain, Member
____
Kellogg O. Lyndes, Member
____
Charles D. Parker, Member

ACKNOWLEDGEMENTS

Thank you to my friends and family - without you I would never have been able to complete this project. You were there when I needed someone to share the exciting parts and when I needed encouragement to continue.

Thank you to my committee, some of which became involved in this investigation when it was nearly done but provided needed insight, and others who showed me that research still has some of the same pitfalls even after the first attempt.

Thank you to Michael Raffin. Without the advice, encouragement and time you provided me, I would never have been able to complete this investigation.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF FIGURES

PAGE

INTRODUCTION

The spondee threshold (ST) is a routine measurement that constitutes part of audiological evaluations (Rupp, 1980). The spondee threshold is used not only as an index of the threshold for speech but also as a method of estimating the accuracy of the pure-tone thresholds.

A variety of methods have been used in the determination of ST. These methods differ according to the psychophysical method (ascending, descending or bracketing) used and increment size (2 or 5 dB) used. There appears to be some disagreement in published research regarding differences in the estimate of ST as a function of psychophysical method. Robinson and Koenigs (1979) reported small but statistically significant differences between ascending and descending ST methods. In addition, Small (1973) reported that errors associated with descending and ascending detection tasks will result in inaccuracy in both methods, while a bracketing method will "cancel" the inaccuracy. In contrast, however, Chaiklin and Ventry (1964) reported that ST is unaffected by the psychophysical method used.

There also appears to exist some disagreement among various researchers concerning the effects of increment size. Discrepancies between 2- and 5-dB increments have been noted by Wilson, Morgan and Dirks (1973). Others (Chaiklin and Ventry, 1964) reported no difference in ST obtained with 2- or 5-dB increments.

There appears to be a general consensus in the published literature that ST may be affected by the actual spondaically stressed words (called spondees for practical purposes throughout this document) used to obtain a ST. Beattie, Edgerton and Svihovec (1975), Beattie, Svihovec, and Edgerton (1975), Bowling and Elpern (1961), Curry and Cox (1966) and Olsen and Matkin (1979) all reported that the Central Institute for the Deaf (CID) spondees have differential intelligibility. The range of intelligibility for the CID spondees is reported to be as great as 10 dB. Through a careful review of the literature, a single study has not been found which controls for relative intelligibility of spondees in a clinical or experimental population during the delivery of the material. The result of Beattie, Forrester and Ruby (1978) would seem to indicate that mode of delivery (recorded vs. monitored live voice) and speaker do not affect the ST.

The ST/pure-tone average (PTA) agreement is used frequently as an index of pure-tone threshold accuracy. The 2- and 3-frequency PTA commonly are used for this purpose (Rupp, 1980). High correlations are reported between these measures in all but some "non-organic" hearing losses. The effect of psychophysical method and increment size on the STPTA agreement has not been investigated systematically in a single study using identical stimuli and procedures.

Other clinical concerns, in addition to accuracy and STPTA agreement, are test-retest reliability and time needed to obtain the ST. Good test-retest reliability is reported for ST obtained via a variety of methods (Chaiklin, Font and Dixon, 1967). Chaiklin and Ventry (1964) reported small differences in the number of spondees needed to obtain an ST. A more detailed review of the literature is contained in Appendix A.

ST accuracy, ST/PTA agreement, test-retest reliability, and number of spondees needed to obtain an ST constitute cost-benefit factors that should be considered when evaluating ST methods. Chial, Beck and VanLandingham (1975) report that "the cost-benefit payoffs of other clinical decision rules are topics deserving study. The results of such investigations should improve our understanding of relative gains and losses of specific procedures" (page 113).

The purpose of the present investigation is to address some of the problems outlined above. Specifically, the present study was designed to determine:

1. ST obtained via several methods,

2. the relationships of the ST to the PTA and

3. the cost effectiveness of ST as a function of methodology employed to obtain the ST. The cost factors are related to test duration, number of words and inaccuracy (discrepancy between the ST and PTA) while benefits are related to test-retest reliabiity and statistical efficiency (Chial *et al*., 1975).

It is proposed that these evaluations will be undertaken by the determination of ST using all possible combinations of three psychophysical methods and two increment sizes. Each such combination will be defined as a test condition.

The null hypotheses associated with the present investigation are:

1. The magnitude of the ST will not be affected by the psychophysical method (ascending, descending or bracketing).

2. The magnitude of the ST will not be affected by the increment size (2 or 5 dB).

3. The number of spondees needed to obtain a ST will not be affected by the test condition used.

4. ST/PTA agreement will not be affected by the test condition used to obtain the ST.

5. Test-retest reliability will not be affected by the test condition.

METHODS

Subjects

The subjects used in the present investigation were 60 adults (younger than 46 years of age) with pure-tone thresholds less than, or equal to, 20-dB HL at the octave frequencies 250 through 8000 Hz. Pure-tone thresholds were obtained using the methods for manual pure-tone threshold audiometry as proposed by the American National Standards Institute (ANSI, 1978). Acoustic-immittance results were commensurate with normal middle-ear transfer function (i.e., in accordance with specifications of guidelines issued by the American Speech, Language and Hearing Association [ASHA, 1979]). Ipsilateral and contralateral acoustic reflexes were present at screening levels (ASHA, 1979). In addition the subjects reported no history of tinnitus, dizziness, otic pain, or any other symptomology consistent with otopathology. For all subjects, the right ear was identified and used as the test ear arbitrarily, unless only the left ear fulfilled the above exigencies. A sample of the consent form obtained from each subject is contained in Appendix B.

Instrumentation

An audiometer (Grasón-Stadler, Model 1701) was used in conjunction with a tape reproducer (Sony TC-377) calibrated in accordance with current standards (ANSI, 1969; ANSI, 1978; NAB, 1965). Calibration data is contained in Appendix C. All testing was accomplished in a sound-treated room (Industrial Acoustic Corporation, Model 403).

Test Materials

The test stimuli consisted of 15 spondees recorded at levels to account for the differential intelligibility reported by Beattie et al. (1975a, 1975b), and Bowling and Elpern (1961), and Curry and Cox (1966). Additional information regarding spondee selection is contained in Appendix D. Spondees were recorded from the Tillman and Olsen spondee recordings contained on the Northwestern University tapes. Re-recording procedures and instrumentation are discussed in Appendix E.

Familiarization

All subjects were familiarized with the test items prior to ST testing. The familiarization procedure consisted of a reading of the test items by the examiner followed by the verbal repetition of each item by the examinee. Any spondee that resulted in an incorrect repetition during this familiarization procedure was eliminated from the list used for the actual ST determinations.

Instructions

Four essential elements were included in the instructions to the subjects:

1. A short description of the nature of the task.

2. Specification of the response expected.

3. Description of the test stimuli.

4. Encouragement to guess when uncertain of the correct response.

The specific instructions are contained in Appendix F.

Test Conditions

ST were determined using three psychophysical methods in conjunction with two increment sizes. Specifically, the following conditions were investigated:

1. Ascending, 2-dB increments and 4 words per level.

2. Ascending, 5-dB increments and 4 words per level.

3. Descending, 2-dB increments and 2 words per level.

4. Descending, 5-dB increments and 5 words per level.

5. Bracketing, 2-dB increments and 2 words per level.

6. Bracketing, 5-dB increments and 5 words per level.

The test conditions were based on a variety of recommendations. The basis for each condition is discussed below.

Ascending

The ascending method was based on the ASHA recommendations (1979). ST determination for each subject began with the presentation of one spondee at the minimum output levels of the audiometer, that is -15-dB HL for the equipment used in this investigation. An additional spondee was presented at 10-dB HL increments (ascending) until a correct verbal response was obtained from the subject. The stimulus level then was reduced by 15 dB, and four spondees were presented. Additional sets of four spondees were presented at 5-dB increments (ascending) until the subject responded correctly to three or more of the spondees at one level. Additional samples began 10 dB below the least-intense level at which three of four spondees were repeated correctly. A minimum of two ascending samples with the same end-point were required for the determination of ST.

The 2-dB ascending method was a modification of ASHA (1979) recommendations and was identical to that described above except that the initial reduction increment was 6 dB rather than 15 dB, and 2-dB ascending steps were used to obtain a ST rather than 5-dB ascending steps. A 6-dB reduction was chosen to parallel the 15 dB recommended by ASHA. Although no explanation was given for the usage of 15 dB, a logical interpretation was assumed to be that the 15 dB represented three times the increment size used. The ST was equal to the lowest level at which 3 of 4 spondees were repeated correctly on at least two ascending samples.

Descending

The descending ST determination used in the present investigation was proposed by Tillman and Olsen (1973). This method entails the presentation of spondees at 10-dB decrements (descending) with one spondee presented at each level. The starting level approximated the patient's comfortable loudness, typically 50-dB HL. When an error was made, another spondee was presented at the same level. If both of the spondees at that level were repeated incorrectly, the level was increased by 10 dB, and the threshold search was begun. If only one spondee was repeated incorrectly, the level was attenuated in 10-dB decrements until two spondees were missed at a given level. When two consecutive errors were encountered, and the presentation level was increased by 10 dB and two spondees were presented at descending 2-dB decrements until five of six responses were incorrect. ST was determined by subtracting the number of correct responses from the starting level then adding one-half the increment used (1 dB for the 2-dB descending method). The correction factor of 1 dB added to the starting level is warranted to maintain the two words/level criterion. The descending method using 5-dB increments parallels the 2 dB procedure, with the exception that 5 spondees were presented at 5-dB decrements and a correction factor of 2 was used based on the recommendation of Wilson, _et al_. (1973).

Bracketing

The bracketing method was based on one advocated by Levitt (1971). The initial spondee was presented at 50-dB HL, and one spondee was presented at 10-dB increments, until an incorrect response was encountered. A total of two spondees were presented at this level. The level then was increased by 2 dB and two more spondees were presented until a level with 100% correct responses was obtained. The intensity was then decreased in 2 dB steps with two spondees at each level until a 0% correct level was obtained. This procedure continued until six changes in direction were noted. The ST was determined as the mean value of the midpoints of the 100% correct and 0% correct levels for the second, fourth and sixth changes in direction. The 5-dB procedure was identical, except that five spondees were presented at 5-dB increments. Appendix G contains flow charts illustrating the test paradigms for each Test Condition.

The subject's ST were determined using each method, and each increment size twice for a total of 12 ST/subject. The order of presentation for the psychophysical methods was counterbalanced and increment size was randomized for the 12 trials. Additional information concerning randomization and counterbalancing is contained in Appendix H.

RESULTS

Spondee threshold values, ST minus the two-frequency PTA, ST minus the three-frequency PTA and the actual number of spondees used to obtain each ST were analyzed using analyses of variance (ANOVA; Ullrich and Pitz, 1981) and the Tukey test (Brunig and Kintz, 1978). Pearson product-moment correlations were obtained of 2- with 3-frequency PTA with the various ST methods used in the present investigation. The 0.01 level of confidence was chosen for these measures. Raw data used for these analyses are contained in Appendix I. Appendix J contains the results of statistical analyses which failed to reveal significant effects.

ST as a Function of Test Condition

Analyses of variance (ANOVA) reveal significant differences In ST values dependent on Psychophysical Method and Increment Size ($p < 0.00000$). Table 1 contains the details of the results of this analysis. Tukey tests were used to determine the relationships between the various increment sizes and psychophysical methods. A comparison of results using 2 dB increments for the three psychophysical methods (ascending, descending and bracketing) reveal statistically significant differences between the ascending method and the other methods

TABLE 1

SPONDEE THRESHOLD ANOVA TABLE

Increment Size (I [2 versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) by Trial (T [trial 1 versus trial 2]) analyses of variance for spondee thresholds.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	45.1501	45.1501	1	13.295	0.0009
Error	183.392	3.3961	54		
P	1469.99	731.493	2	238.206	0.0000
Error	331.650	3.078	108		
IxP	264.143	123.071	2	28.831	0.0000
Error	461.014	4.2687	108		
T	154.290	154.290	1	26.750	0.0000
Error	311.436	5.7673	54		

used with the ascending methods producing an ST which was 1.4 to 2.2 dB greater than either the descending or bracketing method. Descending and bracketing 2 dB methods did not produce statistically different ST. ANOVA and Tukey summary tables are contained in Tables 2 and 3.

Statistically significant differences also were present for the three psychophysical methods when 5 dB increments were used. Again, the ascending methods produced greater ST than either the bracketing or descending method (bracketing 4.7 dB, descending 4.5 dB, $p < 0.0000$). Descending and bracketing 5-dB methods were not significantly different. Results are summarized in Tables 4 and 5.

Ascending ST using 5-dB and 2-dB increments were significantly different with the ascending 5-dB ST producing a ST 2.6 dB greater than the ascending 2-dB method ($p < 0.0000$). Descending and bracketing methods did not produce statistically different ST using 2- or 5-dB increments. A summary is contained in Table 6.

TABLE 2

TWO-dB INCREMENT-SIZE ANOVA
SUMMARY TABLE FOR SPONDEE THRESHOLD

Simple-effects analysis of variance for Psychophysical Method (P [ascending versus descending versus bracketing]) for 2-dB increment spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
P	144.889	72.4447	2	24.66	0.0000
Error	346.697	2.9381	118		

TABLE 3

TUKEY A-POSTERIORI TEST FOR
TWO-dB INCREMENT SPONDEE THRESHOLD

Significance of difference between means for 2-dB increment size as a function of Psychophysical Method (ascending [A2] versus descending [D2] versus bracketing [B2]). Differences significant at the 0.01 level of confidence are indicated by an asterisk.

CONDITION	MEAN	A2	D2	B2
A2	5.8667		1.367*	2.163*
D2	4.4500			0.747
B2	3.7033			

* exceeds the Honestly Significant Difference of 0.929

TABLE 4

FIVE-dB INCREMENT SIZE ANOVA
SUMMARY TABLE FOR SPONDEE THRESHOLD

Simple-effects analysis of variance for Psychophysical Method (P [ascending versus descending versus bracketing]) for 5-dB increment spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
P	853.0420	426.5210	2	90.078	0.0000
Error	558.7320	4.7350	118		

TABLE 5

TUKEY *A-POSTERIORI* TEST FOR 5-DB INCREMENT SPONDEE THRESHOLD.

Significance of differences between means for 5-dB increment size as a function of Psychophysical Method (ascending [A5] versus descending [D5] versus bracketing [B5]). Differences significant at the 0.01 level of confidence are indicated by an asterisk.

CONDITION	MEAN	A5	D5	B5
A5	8.4167		4.733*	4.493*
D5	3.6833			0.24
B5	3.9233			

* exceeds the Honestly Significant Difference of 1.180

TABLE 6

ASCENDING METHOD ANOVA
SUMMARY TABLE FOR SPONDEE THRESHOLD

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for ascending-method spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	195.075	195.0750	1	33.710	0.0000
Error	341.425	5.7869	59		

Number of Spondees

ANOVA and Tukey tests, when appropriate, were used to investigate the effects of psychophysical method, increment size, trial and order of presentation on the number of spondees required to obtain an ST. ANOVA results (Table 7) show an increment size by psychophysical method effect ($p < 0.00000$). Tukey tests were used to investigate this relationship further. The descending 2 dB method required fewer spondees than either ascending or bracketing 2-dB methods (descending = 25.9, bracketing = 41.3, ascending = 40.2; $p < 0.0000$). Ascending 2-dB and bracketing 2-dB methods were not significantly different. The descending 2-dB method required approximately 15 fewer spondees than the ascending 2 dB or bracketing 2 dB methods. A summary of ANOVA and Tukey test results is contained in table 8 through 12.

A 5-dB by psychophysical method comparison of spondees required to obtain the ST reveals that the bracketing 5-dB method uses a significantly greater number of spondees than the ascending 5-dB or descending 5-dB methods (bracketing = 86.7,

TABLE 7

NUMBER OF SPONDEES ANOVA
SUMMARY TABLE

Increment Size (I [2 dB versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) analysis of variance for the number of items required to obtain the spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	39102.3	39102.3	1	429.54	0.0000
Error	4915.76	91.0327	54		
P	168518.0	84259.2	2	879.306	0.0000
Error	10349.1	95.825	108		
IxP	87515.1	43757.6	2	467.073	0.0000
Error	10117.9	93.6846	108		

TABLE 8

TWO-dB INCREMENT SIZE ANOVA
SUMMARY TABLE FOR NUMBER OF SPONDEES

Simple-effects analysis of variance for Psychophysical Method (P [ascending versus descending versus bracketing]) for 2-dB increment for the number of items required to obtain a spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
P	9004.740	4502.37	2	60.88	0.0000
Error	8726.59	73.9541	118		

TABLE 9

ASCENDING METHOD ANOVA
SUMMARY TABLE FOR NUMBER OF SPONDEES

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the ascending Psychophysical Method for the number of items required to obtain a spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	1098.07	1098.07	1	14.47	0.0006
Error	4476.42	75.8716	59		

TABLE 10

BRACKETING-METHOD ANOVA
SUMMARY TABLE FOR NUMBER OF SPONDEES

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the bracketing psychophysical method for the number of items required to obtain a spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	61065.4	61065.4	1	372.19	0.0000
Error	9680.09	164.069	59		

TABLE 11

DESCENDING-METHOD ANOVA
SUMMARY TABLE FOR NUMBER OF SPONDEES

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the descending Psychophysical Method for the number of items required to obtain a spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	549.075	549.07	1	21.21	0.0001
Error	1652.42	28.007	59		

TABLE 12

TUKEY A-POSTERIORI TEST FOR
2-dB INCREMENT SIZE FOR NUMBER OF SPONDEES

Significance of differences between means of number of items required to obtain a spondee threshold for 2-dB increment size as a function of Psychophysical Method (ascending [A2] versus descending [D2] versus bracketing [B2]). Differences significant at the 0.01 level of confidence are indicated by an asterisk.

CONDITION	MEAN	A2	D2	B2
A2	40.1833		14.283*	1.35
D2	25.9000			15.583*
B2	41.5333			

* exceeds the Honestly Significant Difference of 4.663

ascending = 34.1, descending = 30.4; p < 0.0000). The difference in the number of spondees used in the ascending 5 dB and descending 5 dB methods is non-significant (tables 13 and 14).

Increment size by psychophysical method comparisons reveal that, with the exception of ascending methods, 2-dB methods require fewer spondees that 5-dB methods.

The mean values of the six psychophysical method orders had a range of 3.025 spondees (p < 0.62356). Trial 1 and Trial 2 means differed by 1.133 spondees (p < 0.08273).

ST/PTA Agreement

ANOVA and Tukey tests for ST minus the 2-frequency PTA follow the same pattern as the results for ST values. ANOVA (Table 15) reveals a 2-way interaction between increment size and psychophysical method (p < 0.00000). One-way ANOVA and Tukey tests (Tables 16 through 20) reveal that the ascending method using 2- or 5-dB increments produced significantly different ST/PTA agreement than descending or bracketing methods.

TABLE 13

FIVE-dB INCREMENT SIZE ANOVA
SUMMARY TABLE FOR NUMBER OF SPONDEES

Simple-effects analysis of variance for Psychophysical Method (P [ascending versus descending versus bracketing]) for 5-dB increment size for the number of items required to obtain a spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
P	118841.00	59420.0	2	495.26	0.0000
Error	14157.3	119.977	118		

TABLE 14

TUKEY A-POSTERIORI TEST FOR 5-dB
INCREMENT SIZE FOR THE NUMBER OF SPONDEES

Significance of differences between means for 5-dB increment size as a function of Psychophysical Method (ascending [A5] versus descending [D5] versus bracketing [B5]) for the number of spondees required to obtain a spondee threshold. Differences significant at the 0.01 level of confidence are indicated by an asterisk.

CONDITION	MEAN	A5	D5	B5
A5	34.1333		3.783	52.517*
D5	30.3500			53.300*
B5	86.6500			

* exceeds the Honestly Significant Difference of 8.398

TABLE 15

SPONDEE THRESHOLD/TWO-FREQUENCY AVERAGE ANOVA
SUMMARY TABLE

Increment Size (I [2 dB versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) by Trial (T [trial 1 versus trial 2]) analysis of variance for the agreement of spondee threshold with the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	40.3753	40.3753	1	12.284	0.0013
Error	177.488	3.2868	54		
P	1472.43	736.215	2	222.968	0.0000
Error	356.604	3.3019	108		
IxP	207.905	135.452	2	31.420	0.0000
Error	465.588	4.3110	108		
T	152.260	152.260	1	27.128	0.0000
Error	303.078	5.6126	54		

TABLE 16

TWO-dB INCREMENT SIZE ANOVA
SUMMARY TABLE FOR SPONDEE THRESHOLD/TWO-FREQUENCY
PURE-TONE AVERAGE AGREEMENT

Simple-effects analysis of variance for Psychophysical Method (P [ascending versus descending versus bracketing]) for 2-dB increment size for spondee threshold agreement with the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
P	158.117	79.0587	2	26.97	0.0000
Error	345.863	2.9310	118		

TABLE 17

FIVE-dB INCREMENT SIZE ANOVA SUMMARY TABLE
FOR SPONDEE THRESHOLD/TWO-FREQUENCY AVERAGE AGREEMENT

Simple-effects analysis of variance for Psychophysical Method (P [ascending versus descending versus bracketing]) for 5-dB increment size for agreement of spondee threshold with the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
P	885.888	442.944	2	87.90	0.0000
Error	594.638	5.0393	118		

TABLE 18

ASCENDING METHOD ANOVA SUMMARY TABLE FOR
SPONDEE THRESHOLD/TWO-FREQUENCY AGREEMENT

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the ascending Psychophysical Method for the agreement of spondee threshold with the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	183.521	183.521	1	32.54	0.0000
Error	332.799	5.6407	59		

TABLE 19

TUKEY <u>A-POSTERIORI</u> TEST FOR TWO-dB INCREMENT SIZE FOR SPONDEE THRESHOLD/TWO FREQUENCY AGREEMENT

Significance of differences between means for the 2-dB increment size as a function of Psychophysical Method (ascending [A2] versus descending [D2] versus bracketing [B2]) for the agreement of spondee threshold with the two-frequency pure-tone average. Differences significant at the 0.01 level of confidence are indicated by an asterisk.

CONDITION	MEAN	A2	D2	B2
A2	4.455		1.538*	2.245*
D2	2.9167			0.077
B2	2.2100			

* exceeds the Honestly Significant Difference of 0.927

TABLE 20

TUKEY *A-POSTERIORI* TEST FOR 5-dB INCREMENT SIZE FOR SPONDEE THRESHOLD/TWO FREQUENCY AGREEMENT

Significance of differences between means for 5-dB increment size as a function of Psychophysical Method (ascending [A5] versus descending [D5] versus bracketing [B5]) for the agreement of spondee threshold with the two-frequency pure-tone average. Differences significant at the 0.01 level of confidence are indicated by an asterisk.

CONDITION	MEAN	A5	D5	B5
A5	6.9283		4.782*	4.627*
D5	2.1467			0.155
B5	2.3017			

* exceeds the Honestly Significant Difference of 1.217

Descending and bracketing methods produced similar ST/PTA agreement. The ascending 5-dB ST/PTA agreement was significantly poorer than the ascending 2-dB ST/PTA agreement (4.5 versus 6.9 dB; $p < 0.0000$).

ANOVA for ST minus the 3-frequency PTA reveal a complex interaction between Psychophysical Method, Increment Size, Order of presentation and Trial ($p < 0.00932$). All of these dependent variables interact to influence the agreement between the ST and three-frequency PTA. Results of the ANOVA are contained in Table 21. The mean data that yielded this complex interaction may be found in Table 22.

Pearson product-moment correlations (_r_) between each psychophysical method - increment size combination for each trial and the 2- and 3-frequency PTA were computed. The value of _r_ ranges from 0.4128 to 0.6292 for Trial 1 and from 0.4128 to 0.5948 for Trial 2. Using a t-test for differences between dependent correlations the differences in _r_ values for ST minus

TABLE 21

SPONDEE THRESHOLD/THREE FREQUENCY AGREEMENT
ANOVA SUMMARY TABLE

Increment Size (I [2 dB versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) by Trial (T [trial 1 versus trial 2]) by Order of Presentation (O [see Table 22]) analysis of variance for agreement of spondee threshold with the three-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	41.6161	41.6161	1	11.41	0.0017
Error	197.018	3.6485	54		
P	1404.23	702.113	2	199.567	0.0000
Error	379.964	3.5181	108		
IxP	223.225	111.612	2	26.788	0.0000
Error	449.987	4.1665	108		
T	173.952	173.952	1	31.793	0.0000
Error	295.453	5.4714	54		
OxPxT	49.9084	4.991	10	2.543	0.0087
Error	211.923	1.962	108		
OxIxPxT	81.5758	8.1576	10	2.518	0.0093
Error	349.913	3.2399	108		

TABLE 22

DATA FOR SPONDEE THRESHOLD/THREE FREQUENCY AGREEMENT

Spondee threshold minus the three-frequency pure-tone average is shown as a function of test condition, trial, and order. Data are shown for each Psychophysical Method (ascending [A], descending [D], and bracketing [B]) subdivided into the order in which they were presented. The test condition is indicated by the Psychophysical Method followed immediately by the digit representing the increments size (2 dB [2] and 5 dB [5]).

ORDER	TEST CONDITION A2	D2	B2	A5	D5	B5
A-D-B						
TRIAL 1	5.48	3.88	2.45	6.28	4.18	2.96
TRIAL 2	4.24	3.28	2.41	6.78	1.78	2.87
D-A-B						
TRIAL 1	2.37	1.17	1.15	6.47	1.17	1.97
TRIAL 2	2.57	0.47	0.66	3.47	-0.03	0.98
A-B-D						
TRIAL 1	3.61	2.11	2.82	6.91	1.41	3.00
TRIAL 2	3.01	1.61	1.34	5.41	1.61	1.26
D-B-A						
TRIAL 1	1.50	0.30	-1.18	5.00	-0.70	-1.33
TRIAL 2	1.10	-0.20	-1.30	2.50	-1.90	-1.33
B-A-D						
TRIAL 1	2.33	2.13	0.27	5.53	-0.57	-0.47
TRIAL 2	1.53	-0.27	-0.80	0.63	-0.37	-1.13
B-D-A						
TRIAL 1	2.76	-0.04	-0.41	3.16	-0.54	0.26
TRIAL 2	1.36	-1.14	-1.80	3.16	-1.61	-0.42

the 2-frequency PTA versus ST minus the 3-frequency PTA for both trials were not different at the 0.01 level of confidence. Correlation matrices and t-test results are contained in Appendix K and L.

Trial 1 and Trial 2 ST were grouped into the ranges shown in Appendix M and N. For ST minus the 2-frequency PTA, 40% to 81.7% of the ST were within ±5 dB of the PTA. For the ST minus the 3-frequency PTA, 41.7% to 86.7% of the ST were within ±5 dB of the PTA.

Test-Retest Reliability

A comparison of all Trial 1 ST and Trial 2 ST reveal a mean difference of 0.926 dB ($p < 0.00004$) (refer to Table 23). The greatest difference between groups characterized by the order of psychophysical method presentation was 1.695 ($p < 0.93260$).

The difference between mean values for Trial 1 and Trial 2 for the ST/2-frequency PTA was 0.909 dB ($p < 0.00003$). ANOVA summary table is contained in Table 24. The greatest difference in mean values for the six order of presentation of psychophysical method groups was 2.88 dB ($p < 0.05207$).

TABLE 23

RELIABILITY OF SPONDEE THRESHOLD ANOVA
SUMMARY TABLE

Simple-effects analysis of variance for Trial (T [trial 1 versus trial 2]) for spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
T	154.290	154.290	1	26.75	0.0000
Error	311.436	5.7673	54		

TABLE 24

RELIABILITY OF SPONDEE THRESHOLD/TWO FREQUENCY AGREEMENT ANOVA SUMMARY TABLE

Simple-effects analysis of variance for Trial (T [trial 1 versus trial 2]) for the agreement of spondee threshold with the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
T	152.260	152.260	1	27.13	0.0000
Error	303.078	5.6126	54		

Pearson product-moment correlations (r) were computed using each psychophysical method - increment size combination for Trial 1 versus Trial 2. The range of r was from 0.6967 to 0.8861. The poorest r was for ascending 5 dB Trial 1 versus Trial 2 followed by descending 5 dB Trial 1 versus Trial 2. The remaining trial comparisons had a range of 0.08.

A comparison of the ST for Trial 1 versus Trial 2 shows that 98.3% to 100% of Trial 1 ST were within ±5 dB of Trial 2 ST. Results are contained in Appendix O.

Summary

In summary, the null hypotheses that ST would not be affected by psychophysical method and increment size (test condition) is false. In addition, there was an interaction between these two variables. The number of spondees needed to obtain an ST is influenced by the test condition. The ST/PTA agreement, as determined by correlation coefficients, was not influenced by the test condition. Test-retest reliability was not affected by the test condition.

DISCUSSION

The purpose of the present investigation was to determine which Psychophysical Method - Increment Size combination produced the most accurate, most easily replicated and the most cost efficient ST. Two trials using each psychophysical method (ascending, descending and bracketing) and increment size (2- and 5-dB) combination were used to determine 12 ST for each of 60 subjects.

Several of the procedures used in the present investigation were used to minimize subject and experimenter bias. Prior to the collection of data, a set of guidelines for each ST method was developed. These guidelines were followed rigidly during data collection and aided in the elimination of experimenter bias. By counterbalancing psychophysical method and randomizing increment size, effects related to presentation order could be evaluated. By using a set of 15 differentially recorded spondees and familiarizing the subjects before ST were determined, learning effects could be minimized and the differences in spondee intelligibility theoretically were reduced.

ST Values

Analyses of variance (ANOVA) indicated that ascending methods produce ST in poorer agreement with the two-frequency PTA than either 2- or 5-dB descending or bracketing methods. This can be explained by the types of errors that are seen commonly with ascending psychophysical methods. Small (1973) reported that for ascending trials, the subject does not hear the stimulus initially and that an increase in the magnitude of the stimulus may not be detected until the stimulus is at a supra-threshold level. The subject perseverates with the same response until the stimulus is clearly audible (Small, 1973). Therefore, if ascending trials were used exclusively, the result would be an greater ST and poorer agreement between the ST and the PTA. Chaiklin _et al_. (1967) did not find the difference between ascending ST and other ST methods that the present investigation identifies and which psychophysical method error types would predict. Chaiklin _et al_. (1967) reported that the differences between the ST they obtained using an ascending method were "similar to, and in the same direction as...the differences found for STs measured in descending 5 dB (Chaiklin and Ventry, 1964)" (page 143) and the ST were in good agreement with PTA.

Perseverative errors also can be present in descending psychophysical methods. These errors occur when the subject continues to respond when the stimulus is at a level below threshold (Small, 1973). These errors could be assumed not to affect ST determination. The response would be incorrect (except at chance levels) if the subject erroneously perceived the presence of a spondee and would have no effect on the ST. Wilson _et al_. (1973) reported that ST using descending 2- or 5-dB increments are in good agreement and clinically valid. This comment is consistent with the results of the present investigation.

Bracketing and descending methods produced similar ST in the present investigation. Because of the good agreement between these two methods, it can be hypothesized that the ST obtained using either method is an accurate representation of the lowest level at which spondees can be repeated.

The bracketing method is essentially a combination of several ascending and several descending trials and is influenced by the errors associated with each method. A crucial difference in the bracketing method is that levels with 100% correct or 100% incorrect must be obtained before a change in direction (increasing intensity or decreasing intensity) can be made. The slope of the articulation function for spondees can explain, in part, the similarity in ST using descending and bracketing methods. The articulation function for spondees is steepest between 20% and 80% correct (Hudgins, Hawkins, Karlin and Stevens; 1947). At levels less than 20% or greater than 80%, the slope is less steep, and a greater increase in intensity is required for an equal change in percentage correct. By sampling the 100% correct point and the 0% correct point, the errors that are made on psychophysical methods near threshold do not affect the ST.

Number of Spondees

ANOVA for the number of spondees needed for each trial show that the descending 2-dB method required fewer spondees than any other psychophysical method - increment size combination. In clinical practice, one consideration, in addition to accuracy, is the amount of time used for each test. Fewer spondees are required for the descending method and clinical time could be saved by using this method.

In the present investigation, the total number of spondees per level as dictated by the protocol were presented i.e., 2 spondees per level for 2-dB bracketing method, 5 spondees per level for 5-dB bracketing method and 4 spondees for 2- and 5-dB ascending methods. In clinical practice this would not necessarily be the case. For example, in the 5-dB bracketing method 100% of the responses must be correct before the intensity can be decreased or 100% of the responses incorrect before the level can be increased. If a combination of correct and incorrect responses are obtained at a single level, the current direction (increasing or decreasing intensity) would be maintained. In such a case, as few as 2 spondees per level may actually be needed rather than a total of 5 spondees per level. If testing is discontinued at a point where the criteria can no longer be met, fewer spondees and less time may be needed for the determination of ascending or descending ST.

A decrease in the total number of spondees used in either the descending 2- or 5-dB condition would not, however, be possible. In the descending conditions, the total number of spondees correct must be used in the calculation of ST. Spearman (1908, cited in Wilson _et al._, 1973) derived a formula which commonly is used in the determination of threshold. This formula requires not only the total number of correct responses but also the increment used, the initial intensity and the number of possible correct responses per level. If an unequal number of spondees were presented at each level during descending ST determination, the statistical basis for threshold calculation would be violated and an invalid ST estimate would result.

When agreement with the pure-tone results is considered in conjunction with the number of spondees needed for each trial, the descending 2 dB method appears to be the best method in terms of both costs and benefits. It requires the fewest spondees while providing at least as good agreement between the PTA and the ST as the other methods evaluated.

Test-Retest Reliability

While Trial 1 and Trial 2 ST are significantly different ($p < 0.01$), the actual mean difference between trials is less than 1 dB. More than 96% of the two trials were within ±5 dB of each other. This is not inconsistent with the results reported by Chaiklin and Ventry (1964) who reported that the test-retest reliability using 2- or 5-dB increments is within ±6 dB for 93 to 100% of the subjects they evaluated.

ST/PTA Agreement

Pearson product-moment correlations for each ST method, ST minus two-frequency PTA and ST minus three-frequency PTA for trial 1 and 2 yielded correlation coefficients of 0.4128 to 0.6292. These correlations are poorer than those reported in the literature. Chaiklin and Ventry (1964) used descending 2- and 5-dB methods in a similar investigation and they report correlation coefficients ranging from 0.96 to 0.98. Wilson *et al*. (1973) analyzed data gathered on normal and hearing impaired listeners using 2- and 5-dB descending methods and again found correlation coefficients ranging from 0.95 to 0.98. Chaiklin *et al*. (1967) used an ascending 5-dB method and found correlations of 0.96 and 0.97 between the ST and the 2- and 3-frequency PTA. Others (Carhart, 1971; Graham, 1960) have found similar correlations.

A comparison of the methods used in the present investigation with several of the studies reporting high correlations reveal several differences in procedure. Chaiklin *et al*. (1967) used a 5-dB ascending method which consisted of familiarization, and an initial threshold search in ascending 10 dB increments beginning at sub-threshold levels until a correct response occurs. The intensity was then decreased a set amount and ascending sets of 6 spondees were presented at 5 dB increments. Criterion for threshold was 50%. This is essentially the procedure used in the present investigation although 3 of 4 spondees correct was the criterion used for threshold (versus 3 of 6). Chaiklin *et al*. used CID W-1 tapes as the stimuli for ST.

Chaiklin and Ventry (1964) used a descending 2- and 5-dB procedure consisting of familiarization, an initial threshold search, an increase in intensity to supra-threshold levels and a maximum of 6 spondees per level. Threshold was defined as the level at which 50% of the spondees (3 of 6) were repeated correctly. In the present investigation, either 2 or 5 spondees per level were used and threshold was calculated by subtracting the number of spondees repeated correctly from the starting level and subtracting 1 dB from this total for the 2 dB method or 2 dB for the 5 dB method.

Wilson *et al*. (1973) used a procedure identical to the 2- and 5-dB descending procedure used in the present study with one exception. Wilson *et al*. used "36 spondee words...recorded on magnetic tape with the words peaking ($\pm$2 dB) at the level of the 1000 Hz calibration tone," but exactly which words or which recording was used is not reported.

The primary difference between these investigations and the present investigation appears to be the stimuli that were used to obtain the ST. By selectively recording the spondees, the 10-dB range of intelligibility theoretically was reduced so that the spondees should have been approximately equally intelligible at any given intensity level. Thus, the risk of a correct identification due to a spurious recording level was minimized. This change in recording level of the speech material may have changed the relation between the reference threshold level for speech and that for pure tones as promulgated by ANSI standards (ANSI, 1969). Specifically, the reference threshold value for speech recommended by ANSI is based, in part, upon results of studies using spondee materials that were not recorded for equal intelligibility (Jerger, Carhart, Tillman, and Peterson, 1959), and whose results may not be applicable to the current study. It has been proven that, when the Tillman recordings of the spondees are used, the appropriate reference calibration value is changed by about 4 dB, when compared to the reference calibration value for CID materials (Tillman, Johnson, and Olsen, 1966). Thus, some of the discrepancy between ST and PTA observed in the present study, may be the result of the difference between the recording levels of the materials used for the calibration of the audiometer, and the materials and their recording levels as used in the present study. Although the design of the present experiment does not allow for the computation of a correction factor based on the data acquired, it is hypothesized that by reducing the relative intelligibility range, the reference threshold (or correction factor for speech) also would be changed, and that this change would result in a greater ST/PTA agreement.

IMPLICATIONS

If clinical methods are dictated by statistical accuracy and efficiency, results of the present investigation indicate that a descending 2 dB method is optimal for ST determination. This method, as proposed by Tillman and Olsen (1973), begins at supra-threshold levels, descends in 10 dB steps with one spondee at each level until an error is made. The level then is increased by 10 dB and 2 spondees are presented at successive 2 dB decrements until 5 of 6 responses are incorrect. Five of the 6 initial responses in the threshold determination descent must be correct to ensure that a ceiling has been reached. ST is determined by subtracting the number of spondees correct minus one from the starting level.

Future research needs to be conducted to determine whether the relationship between psychophysical method and increment size remain constant when other stimuli are used in ST determination; *i.e.*, using commercially available recordings of CID W-1, Tillman recordings of CID W-1, live voice presentation, *etc*. The effects of the differentially recorded spondees used in the present investigation also need to be further evaluated. Research needs to determine whether the re-recording of the spondees actually reduced the differential intelligibility reported by Bowling and Elpern (1961), Curry and Cox (1966) and Beattie *et al*. (1975a, 1975b). The stimuli used in the present investigation need to be evaluated further to determine an appropriate calibration reference in order to reduce the ST/PTA discrepancies.

REFERENCES

AMERICAN NATIONAL STANDARDS INSTITUTE. Methods for manual pure-tone audiometry (S3.21), 1978.

AMERICAN NATIONAL STANDARDS INSTITUTE. Specifications for audiometers (S3.6-1969[R1973]).

AMERICAN SPEECH, LANGUAGE AND HEARING ASSOCIATION. Guidelines for acoustic immittance screening of middle ear function. ASHA, 1979, 21, 283-289.

AMERICAN SPEECH, LANGUAGE AND HEARING ASSOCIATION. Guidelines for determining threshold for speech. ASHA, 1979, 21, 353-355.

BEATTIE, R.C., EDGERTON, B.J., and SVIHOVEC, D.V. An investigation of Auditec of St. Louis recordings of Central Institute for the Deaf spondees. Journal of the American Auditory Society, 1975a, 1, 97-101.

BEATTIE, R.C., FORRESTER, P.W., and RUBY, B.K. Reliability of the Tillman-Olsen procedure for determination of spondee threshold using recorded and live voice presentations. Journal of the American Auditory Society, 1978, 2, 159-162.

BEATTIE, R.C., SVIHOVEC, D.A., and EDGERTON, B.J. Relative intelligibility of the CID spondees as presented via monitored live voice. Journal of Speech and Hearing Disorders, 1975b, 40, 84-91.

BEATTIE, R.C., SVIHOVEC, D.A., and EDGERTON, B.J. Comparison of speech detection and spondee thresholds and half- vs. full-list intelligibility scores with MLV and taped presentations of NU#6. Journal of the American Auditory Society, 1978, 3, 276-272.

BERGER, K. Speech audiometry. In D.E. Rose (Ed.), Audiological assessment. Englewood Cliffs: Prentice-Hall, 1971.

BODE, D.L., and CARHART, R. Stability and accuracy of adaptive tests of speech discrimination. Journal of the Acoustical Society of America, 1974, 56, 963-970.

BODE, D.L., and CARHART, R. Estimating CNC discrimination with spondee words. Journal of the Acoustical Society of America, 1975, 57, 1216-1218.

BOWLING, L.S., and ELPERN, B.S. Relative intelligibility of items on CID Auditory Test W-1. Journal of Auditory Research, 1961, 1, 152-156.

BRUNIG, J.L. AND KINTZ, B.L. Computational handbook of statistics (2nd. ed.). Glenview, Illinois: Scott, Foresman and Company, 1978.

BURKE, L.E. and NERBONNE, M.A. Influence of the guess factor on the speech reception threshold. Journal of the American Auditory Society, 1978, 4, 87-90.

CARHART, R. Observations on relations between thresholds for pure-tones and for speech. Journal of Speech and Hearing Disorders, 1971, 36, 476-483.

CARHART, R., and PORTER, L.S. Audiometric configuration and prediction of thresholds for spondees. Journal of Speech and Hearing Research, 1971, 14, 486-495.

CHAIKLIN, J.B. The relation among three selected auditory speech thresholds. Journal of Speech and Hearing Research, 1959, 2, 237-243.

CHAIKLIN, J.B., FONT, J., and DIXON, R.F. Spondee thresholds measured in ascending 5 dB steps. Journal of Speech and Hearing Research, 1967, 10, 141-145.

CHAIKLIN, J.B., and VENTRY, I.M. Spondee threshold measurments: a comparison of 2- and 5-dB methods. Journal of Speech and Hearing Disorders, 1964, 29, 47-59.

CHIAL, M.R., BECK, W.G., and VANLANDINGHAM, G. Cost-benefit analysis of two dicision criteria for ST measurements. Journal of Speech and Hearing Research, 1975, 18, 105-114.

CONN, M., DANCER, J., and VENTRY, I.M. A spondee list for determining speech reception threshold without prior familiarization. Journal of Speech and Hearing Disorders, 1975, 40, 388-396.

CURRY, E.T., and COX, B.P. The relative intelligibility of spondees. Journal of Auditory Research, 1966, 6, 419-424.

GRAHAM, J.T. Evaluation of methods for predicting speech reception threshold. Archives of Otolaryngology, 1960, 72, 347-350.

HAGERMAN, B. Reliability in the determination of speech reception threshold (SRT). Scandinavian Audiology, 1979, 8, 195-202.

HIRSH, I.J., DAVIS, H., SILVERMAN, S.R., REYNOLD, E.G., ELDERT, E., and BENSON, R.W. Development of materials for speech audiometry. Journal of Speech and Hearing Disorders, 1952, 17, 321-337.

HODGSON, W.R. Basic audiologic assessment. Baltimore: Williams and Wilkins, 1980.

HOPKINSON, N.T. Speech reception threshold. In J. Katz (Ed.), Handbook of clinical audiology (2nd ed.). Baltimore: Williams and Wilkins, 1978.

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

JERGER, J.F., CARHART, R., TILLMAN, T.W., and PETERSON, J.L. Some relations between normal hearing for pure-tones and for speech. Journal of Speech and Hearing Research, 1959, 2, 126-140.

LEVITT, H. Transformed up-down methods in psychoacoustics. Journal of the Acoustical Society of America, 1971, 49, 467-477.

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

MARTIN, F.N., and STAUFFER, M.L. A modification of the Tillman-Olsen method for obtaining speech reception thresholds. Journal of Speech and Hearing Disorders, 1975, 40, 25-28.

OLSEN, W.O., and MATKIN, N.D. Speech audiometry. In W.F. Rintlemann (Ed.), Hearing assessment. Baltimore, MD: University Park Press, 1979.

NATIONAL ASSOCIATION OF BROADCASTERS. Standard magnetic tape recording and reproducing (reel-to-reel). Washington, D.C., 1965.

ROBINSON, D.O., and KOENIGS, M.J. A comparison of procedures and materials for speech reception thresholds. Journal of the American Auditory Society, 1979, 227-330.

RUPP, R.R. Classical approaches to the determination of the spondee threshold. In R.R. Rupp, & K.G. Stockdell (Eds.), Speech protocols in audiology, New York: Grune & Stratton, 1980.

SIEGENTHALER, B.M., and STRAND, R. Audiogram average methods and SRT scores. Journal of the Acoustical Society of America, 1964, 36, 589-593.

SMALL, A.M. Psychoacoustics. In F.D. Minifie, T.J. Hixon, & F. Williams (Eds.), Normal aspects of speech, hearing and language. Englewood Cliffs: Prentice-Hall, 1973.

THORNTON, A.R., and RAFFIN, M.J.M. Speech discrimination scores modeled as a binomial variable. Journal of speech and hearing research, 1978, 21, 507-518.

TILLMAN, T.W., and JERGER, J.F. Some factors affecting the spondee threshold in normal-hearing subjects. Journal of Speech and Hearing Research, 1959, 2, 145-146.

TILLMAN, T.W., and OLSEM W.O. Speech audiometry. In J.F. Jerger (Ed.), Modern developments in audiology (2d. Ed.). New York: Academic Press, 1973.

TILLMAN, T.W., JOHNSON, R.M., and OLSEN, W.O. Earphone versus sound-field threshold sound-pressure levels for spondee words. Journal of the Acoustical Society of America, 1966, 39, 125-140.

ULLRICH, J.F., and PITZ, G. A general purpose analysis of variance routine. Unpublished manuscript, University Computer Center, University of Montana, Missoula, Montana, 1981.

WILSON, R.H., and CARHART, R. Influence of pulsed masking on the threshold for spondees. Journal of the Acoustical Society of America, 1969, 46, 998-1010.

WILSON, R.H., MORGAN, D.E., and DIRKS, D.D. A proposed SRT procedure and its statistical precedent. Journal of Speech and Hearing Disorders, 1973, 38, 184-191.

YOST, W.A., and NIELSEN, D.W. Fundamentals of hearing. New York: Holt, Rinehart, & Winston, 1977.

APPENDIX A

REVIEW OF THE LITERATURE

Speech stimuli are used routinely in assessing auditory sensitivity. One commonly used measure is the spondee threshold (ST). In the determination of ST, two-syllable words are spoken with equal stress on each syllable to determine the threshold for speech. Also known as the spondaic-word threshold, speech-reception threshold, threshold for spondaic words, and spondaic threshold (Hopkinson, 1978), the ST has several purposes. According to Rupp (1980), it establishes a hearing level for a carefully defined speech signal which in turn can aid in the estimation of communicative handicap. ST provides verification of pure-tone results and provides a basis for setting the level used in speech discrimination testing. A comparison between aided and unaided ST may be used to assess the benefits of amplification.

History

Test materials used to obtain ST originated during World War II at the Harvard Psychoacoustic Laboratories (PAL) by Hudgins, Hawkins, Karlin and Stevens (1947). Auditory Test #9 and #14 were developed to assess hearing for speech and consisted of 84 spondaically stressed words. Auditory Test #9 was recorded at a constant level while #14 decreased in intensity at the rate of 4 dB per group of four spondees. Criteria for the selection of spondees were phonetic dissimilarity, phonetic composition approximating English, homogeneous intensity and intelligibility, and words familiar to adults who would be tested using these materials. All words consisted of two syllables with equal stress on both syllables. These spondees then were randomized to form PAL Auditory tests #9 and #14.

The first major revision of PAL Auditory tests #9 and #14 was made by Hirsh, Davis, Silverman, Reynolds, Eldert and Benson (1952). The list of 84 spondees was evaluated and the 36 most familiar words were used to form Central Institute for the Deaf Auditory Test W-1 (CID W-1). Six randomizations of these 36 words were recorded at a constant level with a carrier phrase "Say the word" recorded at a level 10 dB greater than that of the spondee. These recordings subsequently were presented to a group of normal listeners and some words were judged to be more difficult or easier than others. Difficult words were increased by 2 dB and easy words were decreased by 2 dB in a second set of recordings. Hirsh, <u>et al</u>. (1952), reported that this modification resulted in more homogeneous stimulus items and less variation in patient responses.

CID Auditory Test W-2 was subsequently developed and consists of the same set of spondees as CID W-1. The CID W-2 format has sets of three spondees recorded at a specific level with subsequent sets of three words 3 dB weaker than the previous set (Rupp, 1980; Olsen and Matkin, 1979; Hopkinson, 1978; Hirsh, et al., 1952). CID W-1 is reported to be the most commonly used test for ST followed by CID W-2. These comprise 72.3% of the responses obtained in a survey by Martin and Pennington (1971). CID W-1 is used most commonly by 57.9% and CID W-2 is used most commonly by 14.4% of the 276 audiologists responding.

Paradigm for the Determination of Spondee Threshold

There are four distinct components in the determination of ST. The first phase is instruction, followed by familiarization, orientation, and the actual determination of the ST (Rupp, 1980).

Instructions

As reported in ASHA (1979), instruction for ST measurement ideally contains four elements: orienting to the task, specifying of the expected response, informing the subject of the nature of the stimuli (speech) and encouraging guessing when the subject is unsure of the correct response. Each of these elements serves a function that results in a more rapid and accurate estimate of the ST. The first three elements generally inform the subject about the nature of the ST while the fourth element, encouraging guessing, resulted in a more accurate estimate of ST. Burke and Nerbonne (1978) reported a mean difference of 4.2 dB in the ST of subjects instructed to guess if they were unsure versus those who were not instructed to guess.

Familiarization

Familiarization also dramatically influences the ST. Jerger and Tillman (1959) reported that ST is as much as 4 to 5 dB more sensitive and more stable when the subject is familiarized with the test words before determining the ST than when there is no familiarization. Conn, Dancer and Ventry (1975) presented 18 of the CID W-1 words at the subject's ST level with familiarization and the other 18 words without familiarization. No attempt was made to group the spondees according to relative intelligibility. Familiarization consisted of the experimenter reading the spondees in alphabetical order and having the subject repeat the words. The differences between correct, incorrect, and no response were statistically significant for the two conditions and reaffirm the need for familiarization if the entire list of 36 spondees is used.

A variety of familiarization procedures are reported in the literature. Some procedures (Conn, et al., 1975) require the subject to repeat spondees spoken by the tester. Others require the subject to read a printed set of spondees to the examiner (Martin and Stauffer, 1975) or, the examiner to read a list of spondees to the subject (Wilson and Carhart, 1969). Some require the subject to read the spondees silently either from an alphabetical list (Hopkinson, 1978) or from index cards each with one spondee printed on it (Chaiklin, 1959).

Orientation

The orientation phase of ST determination is the initial search for the approximate ST. It begins at supra-threshold levels and rapidly decreases in intensity to near threshold. This prepares the subject for the threshold determination task. In ascending methods, which begin below thershold, an orientation phase is not used in determination of ST (Rupp, 1980).

Threshold Determination

In the threshold determination phase, the ST in dB HL is found for each test ear. A variety of factors including psychophysical method, increment size and stimulus words used will influence the ST.

Psychophysical method: The methods used in the determination of ST are based on classical psychophysical methods. These psychophysical methods include the method of adjustment and the method of limits. The essential components of the method of adjustment are a series of ascending and descending trials with a countinuously variable increment size along the test stimulus parameter, and direct listener control of the stimulus. The method of limits also utilizes ascending and descending trials but the test parameter varies in set increments rather than on an infinitely variable scale. The experimenter has control over the stimulus presentation level but is influenced by the subject's response (Small, 1973; Yost and Nielsen, 1977).

Determination of the ST by Békésy tracking is similar to the method of adjustment while other methods more closely resemble the method of limits. A simple up-down adaptive procedure in which the experimenter increased the signal by a set increment following an incorrect response and decreased the signal by the same amount following a correct response has been used in the determination of ST. Threshold is determined to be the center point around which the correct and incorrect responses are clustered. Other methods utilizing only ascending or descending trials are related to traditional methods in that the experimenter varies the stimulus in set increments but violates the traditional procedures in requiring only ascending or descending trials (Small, 1973).

Three general methods mentioned above, up-down (bracketing), ascending and descending, are used clinically. Martin and Pennington (1971) reported that the most frequently used is bracketing (43.8%) followed by descending (35.9%) and ascending methods (20.3%).

A bracketing or simple up-down method has been used by a variety of researchers including Bode and Carhart (1974, 1975), Chial, et al. (1975), and Hagerman (1979). They reported that bracketing is a valid and reliable method of determining ST.

Descending methods have been used clinically and in research by Bode and Carhart (1975), Chaiklin (1959), Conn, et al. (1975), Robinson and Koenig (1979), Tillman and Olsen (1973) and Wilson, et al. (1973). Descending methods begin at supra-threshold levels and initially use large increment steps to get near threshold. After several errors are made, the level usually is increased by a set amount and the descending threshold determination is begun. A constant number of words presented at each predetermined level until a termination criterion - usually in percent or a certain number of errors - is met. Threshold is determined by subtracting the number correct minus one from the starting level (Tillman and Olsen , 1973) or as the lowest level meeting the 50% criterion for threshold (Rupp, 1980) or 2 dB above the 0% correct level (Hopkinson, 1978).

An ascending ST determination is recommended by ASHA (1979) and is used extensively in the published literature. Numerous researchers have used ascending methods to evaluate the relative intelligibility of spondees or the relationship between ST and other audiometric tests (Bowling and Elpern, 1961; Chaiklin and Ventry, 1959; Chaiklin, et al., 1967; Chaiklin and Ventry, 1964; Conn, et al., 1975; Curry and Cox, 1966). ASHA (1979), and Chaiklin et al. (1967) exemplify two of the few research projects that advocate the clinical use of an ascending ST method as a routine procedure. Others (Conn, et al., 1975; Rupp, 1980) report that the primary value of ascending methods is their application in the delineation of non-organic hearing impairment. These researchers found a more precise PTA/ST agreement for ascending methods when ascending and descending results were compared in the evaluation of non-organic hearing impairment.

Robinson and Koenigs (1979) found that ascending methods produced statistically significant differences in the estimate of the ST (the average difference being 1.22 dB), but concluded that this difference was not clinically relevant and, therefore, that either ascending or descending methods produced valid and reliable ST.

<u>Increment size</u>: An additional variable influencing the ST is the increment size used during threshold determination. The most commonly used increments are 2 dB and 5 dB, although some researchers have used 4-dB increments. Chaiklin and Ventry (1964) compared the ST obtained using a descending method in 5-dB and 2-dB steps. They reported no statistically significant differences in the ST using either increment. Test-retest reliability was asserted to be "good" with both methods: 93% within ±6 dB for 2-dB steps (none greater than ±8%), and 100% within ±5 dB for 5-dB steps.

Although Chaiklin and Ventry (1964) reported no statistically significant differences between 2-dB and 5-dB methods, the slope of the performance-intensity function for spondees predicted that there would be a significant difference between methods using 2-dB and 5-dB increments. The slope of the performance-intensity function between 20% and 80% is 8 to 12%/dB

(Hudgins, et al., 1947; Beattie, et al., 1975b). In a method attempting to estimate 50% intelligibility, a 1-dB increase in the presentation level can increase the performance score by 16% (e.g.: 50% to 66%, or 42% to 58%). This dramatic increase within a small range of intensity change may lead to the inference that greater precision could be obtained theoretically through the usage of a smaller increment size during the determination of ST (Hopkinson, 1978). Wilson, et al. (1973), compared 2-dB and 5-dB steps for descending methods with two words per level in the 2-dB method, and five words per level in the 5-dB method. They found a statistically significant ($p < 0.01$) difference of 1.2 dB between these two increment sizes, but they concluded that this was not a clinically meaningful difference.

Materials: Characteristics of the spondees used may also influence the ST. One of these characteristics is intelligibility. One of the criteria for inclusion of words as test material used by Hudgins, et al. (1947) and other researchers was homogeneous intelligibility. Bowling and Elpern (1961) evaluated the relative intelligibility of the CID W-1 spondees and found an intelligibility range of 10 dB. The "easiest" (the most readily correctly identified at reduced presentation levels) spondee (workshop) was identified at 2.7-dB HL. The 22 spondees in the center of the range of intelligibility varied by only 3.5 dB, and are recommended for use in the ST determination. Curry and Cox (1966), in a similar study, found an intelligibility range of 8.1 dB and suggested that the items at the extrema of the intelligibility range be eliminated to produce a 27-spondee list with a range of 4 dB. Beattie, *et al.* (1975a), proposed an 18-spondee list with an intelligibility range of 1.5 dB for the determination of ST.

A summary of six studies evaluating the relative intelligibility of spondees was reported by Olsen and Matkin (1979). Their results suggested that all six studies agreed on the homogeneity of 4 words, while there was less agreement for 15 homogeneous words. On that basis, Olsen and Matkin (1979) recommended that a set of 15 or 24 spondees (with greater consensus of homogeneity) be used rather than the 36 CID spondees.

In a related study, Conn, *et al.* (1975) evaluated the relative familiarity of the 36 CID spondees and concluded that 17 of them could be used reliably without familiarization. These spondees also are among the group of spondees reported to be more homogeneous by Olsen and Matkin (1979).

Spondees per level: An additional variable found in ST measurement is the number of spondees presented at each level. Hagerman (1979) reported that the clinical use of 10 trials/level results in an accurate and reliable ST. Computer simulation of results obtained with 10, 5 and 4 trials/level (using a 5-dB increment, bracketing method), however, indicated that 10 trials/level is not optimal. Results obtained with 4 trials/level are more reliable than those obtained with 10 trials/level, and the time saved was reportedly 9%. Five trials/level was the most reliable, but the time saved was only 3% over the 10 trials/level method. Other researchers (e.g., Hodgson, 1980) reported that "...there is a tendency for speech threshold to get better as fewer words are used to determine threshold. If only 3 or 4 words are used, the speech threshold may be quite close to the threshold of detectability." Not only was this assertion not documented with any data, but it clearly contradicts a basic mathematical and statistical theorem that indicates that the error of the measurement is inversely proportional to the number of trials used for the measurement (Thornton and Raffin, 1978).

Relationship to Other Measures

One of the purposes of ST as reported by Rupp (1980) is as a consistency check with the PTA. A variety of researchers have discussed the relationship between pure-tone tests and speech tests. Beattie, et al. (1975a), reported that the ST is approximately 7.6 dB greater than the speech-awareness or the speech-detection threshold. Rupp (1980) reported the difference between these two measurements as being 12 dB, while Chaiklin (1959) asserted that the difference is 9 dB. Bode and Carhart (1975) obtained speech-discrimination scores (using monosyllabic words) at the level of the ST. The speech-discrimination score obtained at this level was between 23% and 30%.

Siegenthaler and Strand (1964) evaluated seven methods for predicting ST from the pure-tone test results. These methods include two-frequency PTA , three-frequency PTA (500, 1000, 2000 Hz), two regression equations, the American Medical Association percent hearing-loss method, and a weighted mean of pure-tone thresholds from 250 through 4000 Hz. These researchers concluded that the two-frequency PTA (for the most sensitive thresholds between 500 and 2000 Hz) is the best single predictor of ST regardless of audiometric configuration. Tillman and Olsen (1973) also agreed that the two-frequency PTA is the best predictor of ST. Carhart (1971), and Carhart and Porter (1971) proposed that the thresholds at 500 Hz and 1000 Hz be averaged and that a 2-dB correction factor be subtracted to predict most accurately the ST, and to optimally limit the deleterious effects of hearing-loss configuration on this prediction. Some researchers have concluded that either a two-frequency or a three-frequency PTA will be adequate predictors of ST (Chaiklin, *et al.*, 1967; Graham, 1960; Hopkinson, 1978; Wilson, *et al.*, 1973). All of these investigators cited correlations between the two PTA methods and the ST equal to, or greater than, 0.95. In addition, they also have tended to report that if the PTA/ST agreement is not within ±6 to ±10 dB, additional testing is indicated to rule out the existence of non-organic hearing impairment, and tester and equipment malfunctions that may affect the ST adversely.

Summary

ST has been used routinely in audiometric test batteries. A variety of procedures have been used for the determination of ST including ascending, descending, and bracketing methods with 2-dB or 5-dB increments. The effects of methodologic differences (*e.g.*: psychophysical method, increment size) on ST and its relationship to pure-tone thresholds have not been investigated systematically in any single study. In summary, a variety of factors influence the ST. These include instructional set, familiarization, psychophysical method, increment size and spondees used.

APPENDIX B

CONSENT FORM

I hereby agree to act as a volunteer listener for the purposes of gathering data for a Master´s Thesis. I understand that my name or other identifying information will not be used in the research. I also understand that there is no physical risk or discomfort associated with the tests being used. The test procedures have been explained to me and I understand how I am to be involved in this project.

Subject´s signature _______________.

Date _______________.

APPENDIX C

TAPE RECORDER/REPRODUCER CALIBRATION

Three tape recorders/reproducers were used during the course of the present investigation. One instrument (Akai, GX-635D) was used as a recording unit for the dubbing of the test spondees from the Tillman recordings. Another instrument (Akai, 1722W) was used as the reproducing unit for the Tillman recordings during the dubbing procedures. The last instrument (Sony, TC-377) was used as the reproducing unit during the testing of subjects, and is the unit associated with the test facilities at the Speech, Hearing and Language Clinic.

Tape heads of the three tape recorder/reproducers used in the present investigation were cleaned and degaused prior to calibration and on a regular basis during re-recording and data collection.

<u>Tape Speed</u>

The tape-drive speed of each tape recorder/reproducer system was evaluated using a 5-minute timing tape and a stop watch. The tape speeds for the Akai 1733W, Akai GX-635D and Sony TC-377 were all within 0.2% of the nominal value of the tape. This is in accordance with specifications promulgated by the National Association of Broadcasters (NAB, 1965).

Playback/Record-Level Calibration

The playback mode of the Akai 1722W and the Sony TC-377 tape recorder/reproducers was evaluated using a commercial reproduce alignment tape (Ampex #01-31321-01) with 50 us equalization and a tape speed of 19.01 cm/s. The output of the tape recorder/reproducers were monitored via a vacuum-tube voltmeter (VTVM) (Hewlett-Packard, Model 400HR). Output levels for the Akai 1722W (input to the Akai GX-635D) were within NAB standard reproducing system response-limits (1965). Output levels of the Sony TC-377 were within NAB reproducing and recorded-response limits for special purpose systems (1965).

The frequency response of the record mode of the Akai GX-635D was evaluated using an audio-frequency spectrometer (Bruel and Kjaer [B and K] Type 2112) and a graphic-level recorder (B and K, Type 2305). Values obtained were within the levels promulgated by NAB standard recorded response limits (1965). Calibration data are contained in Tables C1, C2, and C3.

TABLE C1

TAPE RECORDER/REPRODUCER CALIBRATION
Akai GX-635D

NOMINAL FREQUENCY (Hz)	CHANNEL 1	CHANNEL 2	TOLERANCE RE: NAB STANDARDS (relative dB)
700	0	0	--------
15000	-0.7	-1.5	+1 to -3
12000	-0.2	+0.1	+1 to -2
10000	+0.4	+0.4	±1
7500	+0.6	+0.3	±1
5000	0.0	+1.0	±1
2500	+0.9	+0.3	±1
1000	-0.1	-0.7	±1
500	+0.2	-0.1	±1
250	+0.9	+0.8	±1
100	-0.2	+0.6	±1
50	-1.0	-1.2	+1 to -3

TABLE C2

TAPE RECORDER/REPRODUCER CALIBRATION
Sony TC-377

NOMINAL FREQUENCY (Hz)	CHANNEL 1	CHANNEL 2	TOLERANCE RE: NAB STANDARDS (relative dB)
700	0	0	--------
15000	-4.0	-5.0	--------
12000	-4.0	-5.0	--------
10000	-3.0	-4.5	--------
7500	-3.0	-4.0	-5 to +2
5000	-2.0	-4.5	±2
2500	-1.4	-1.8	±2
1000	-0.4	-0.3	±2
500	+0.3	+0.2	±2
250	+1.3	+1.2	±2
100	-2.5	+2.2	-5 to +2
50	+0.5	+1.0	--------

TABLE C3

TAPE RECORDER/REPRODUCER CALIBRATION
Akai 1722W

NOMINAL FREQUENCY (Hz)	CHANNEL 1	CHANNEL 2	TOLERANCE RE: NAB STANDARDS (relative dB)
700	0	0	--------
15000	-4.0	----	+1 to -3
12000	-2.0	----	+1 to -2
10000	-2.0	----	±1
7500	-1.0	----	±1
5000	-0.5	----	±1
2500	0.0	----	±1
1000	0.0	----	±1
500	+1.0	----	±1
250	+1.5	----	±1
100	+2.0	----	±1
50	0.0	----	+1 to -3

APPENDIX D

SELECTION OF SPONDEES

The 15 spondees used in the present investigation were chosen to reflect homogeneous intelligibility. Beattie *et al.* (1975a; 1975b), Bowling and Elpern (1961), and Curry and Cox (1966) evaluated the intelligibility of the CID spondees. Each developed a relative intelligibility scale based on the level at which the subjects in their studies correctly identified the spondee. The mean relative-dB value and the range for each spondee were determined for these studies. Spondees with a range no greater than 4 dB then were evaluated. From the list of 17 spondees with a range no greater than 4 dB, two spondees (*duckpond* and *eardrum*) were eliminated because of the large standard deviation (greater than 5.0 dB) for these spondees reported by Beattie *et al.* (1975a). The remaining 15 spondees were used in the present investigation. Table D1 contains the relative-dB levels, means and ranges used for spondee selection.

TABLE D1

SPONDEE SELECTION

Relative level of each spondee words to obtain equal intelligibility reported by several investigators: Beattie, *et al.* (1975a) [A]; Beattie, *et al.* (1975B) [B]; Bowling, & Elpern (1961) [C]; Curry, & Cox (1966) [D]. The mean level for equal intelligibility from these four studies is indicated (mean), as is the range of means from these studies. The absolute value of the range of reported equal intelligibility is indicated in the column heading of **. Single asterisks preceding the spondees indicate those items used in the present study.

SPONDEES	A	B	C	D	MEAN	RANGE	**
*airplane	4.7	7.5	6.7	4.2	5.8	4.2 - 7.5	3.3
*armchair	7.5	7.5	6.6	4.4	6.5	4.4 - 7.5	3.1
baseball	3.0	7.7	6.1	3.0	5.0	3.0 - 7.7	4.7
birthday	6.3	11.2	8.7	5.6	8.0	5.6 - 11.2	5.6
cowboy	4.3	10.0	5.8	5.0	6.3	4.3 - 10.0	5.7
*daybreak	8.3	9.2	9.3	6.6	8.4	6.6 - 9.2	2.6
doormat	8.3	12.1	9.0	6.8	9.1	6.8 - 12.1	5.3
drawbridge	8.2	14.1	7.9	4.8	8.6	4.8 - 14.1	9.3
duckpond	8.3	11.8	8.8	10.1	9.8	8.3 - 11.8	3.5
eardrum	6.9	9.4	8.5	6.1	7.7	6.1 - 9.4	3.3
farewell	6.3	11.0	12.3	8.0	9.4	6.3 - 12.3	6.0
grandson	12.2	12.9	9.7	8.4	10.8	8.4 - 12.9	4.5
greyhound	7.0	12.6	8.4	6.0	8.5	6.0 - 12.6	6.6

TABLE D1 (continued)

SPONDEES	A	B	C	D	MEAN	RANGE	**
hardware	3.5	9.6	7.5	4.9	6.4	3.5 - 9.6	6.1
*headlight	8.3	11.3	9.8	8.6	9.5	8.3 - 11.3	3.0
*horseshoe	7.8	11.1	8.6	8.6	9.0	7.8 - 11.1	3.3
hotdog	5.0	8.6	4.4	3.8	5.5	3.8 - 8.6	4.8
hothouse	12.7	11.7	7.6	10.9	10.8	7.6 - 12.7	5.1
iceberg	4.8	10.5	7.3	4.0	6.7	4.0 - 10.5	6.5
*inkwell	8.8	9.5	7.2	7.0	8.1	7.0 - 9.5	2.5
*mousetrap	7.7	9.4	7.8	7.7	8.0	7.7 - 9.4	1.7
*mushroom	10.3	10.6	7.0	7.7	8.9	7.0 - 10.6	3.6
northwest	5.2	10.5	8.6	6.2	7.6	5.2 - 10.5	5.3
oatmeal	7.0	12.1	8.6	1.6	8.8	1.6 - 12.1	10.5
*padlock	8.3	10.8	8.7	8.0	8.9	8.3 - 10.8	2.5
pancake	7.7	16.5	9.9	8.0	10.5	7.7 - 16.5	8.8
playground	4.5	11.5	7.4	4.6	7.0	4.5 - 11.5	7.0
railroad	4.8	10.7	8.4	6.5	7.6	4.8 - 10.7	11.9
*schoolboy	8.5	10.6	9.5	6.9	8.9	6.9 - 10.6	3.5

TABLE D1 (continued)

SPONDEES	A	B	C	D	MEAN	RANGE	**
*sidewalk	6.0	8.1	6.8	6.5	6.9	6.0 - 8.1	2.1
*stairway	6.3	9.7	8.4	8.0	8.1	6.3 - 9.7	3.4
*sunset	6.2	9.9	9.3	6.2	7.9	6.2 - 9.9	3.7
*toothbrush	7.7	10.2	9.0	8.0	8.7	7.7 - 10.2	2.5
*whitewash	7.2	9.5	7.4	7.8	8.0	7.2 - 9.5	2.3
woodwork	4.3	14.9	8.2	4.8	8.0	4.3 - 14.9	10.6
workshop	2.7	10.9	6.2	2.8	5.7	2.7 - 10.9	8.2

APPENDIX E

RECORDING OF SPONDEES

Spondees were recorded from the Tillman-Olsen recordings of Northwestern University Auditory Tests using an Akai 1722W and an Akai GX-635D tape recorder/reproducer in conjunction with an audio-frequency spectrometer (B and K, Type 2112) and a graphic-level recorder (B and K, Type 2305). A block diagram of the instrumentation used to record the spondees is contained in Figure E1.

The 15 spondees were recorded so that syllable peaks were within approximately 1 dB of each other. Recording levels were based on the mean values as determined from the studies by Beattie _et_ _al_. (1975a; 1975b), Bowling and Elpern (1961) and Curry and Cox (1966). All spondees were recorded relative to the level of the highest mean dB level. Thus, _headlight_, with a mean value of 9.5 dB, peaked at 0 VU (relative to the 1000 Hz calibration tone). _Airplane_ peaked near -3.7 dB (mean value = 5.8 dB). Mean dB values and recording levels are contained in Table E1. Hard-copy tracings of the spondees, as measured from Channel 1 and 2, are contained in Figure E2.

FIGURE E1

Block diagram of the instrumentation used for the recording of stimuli.

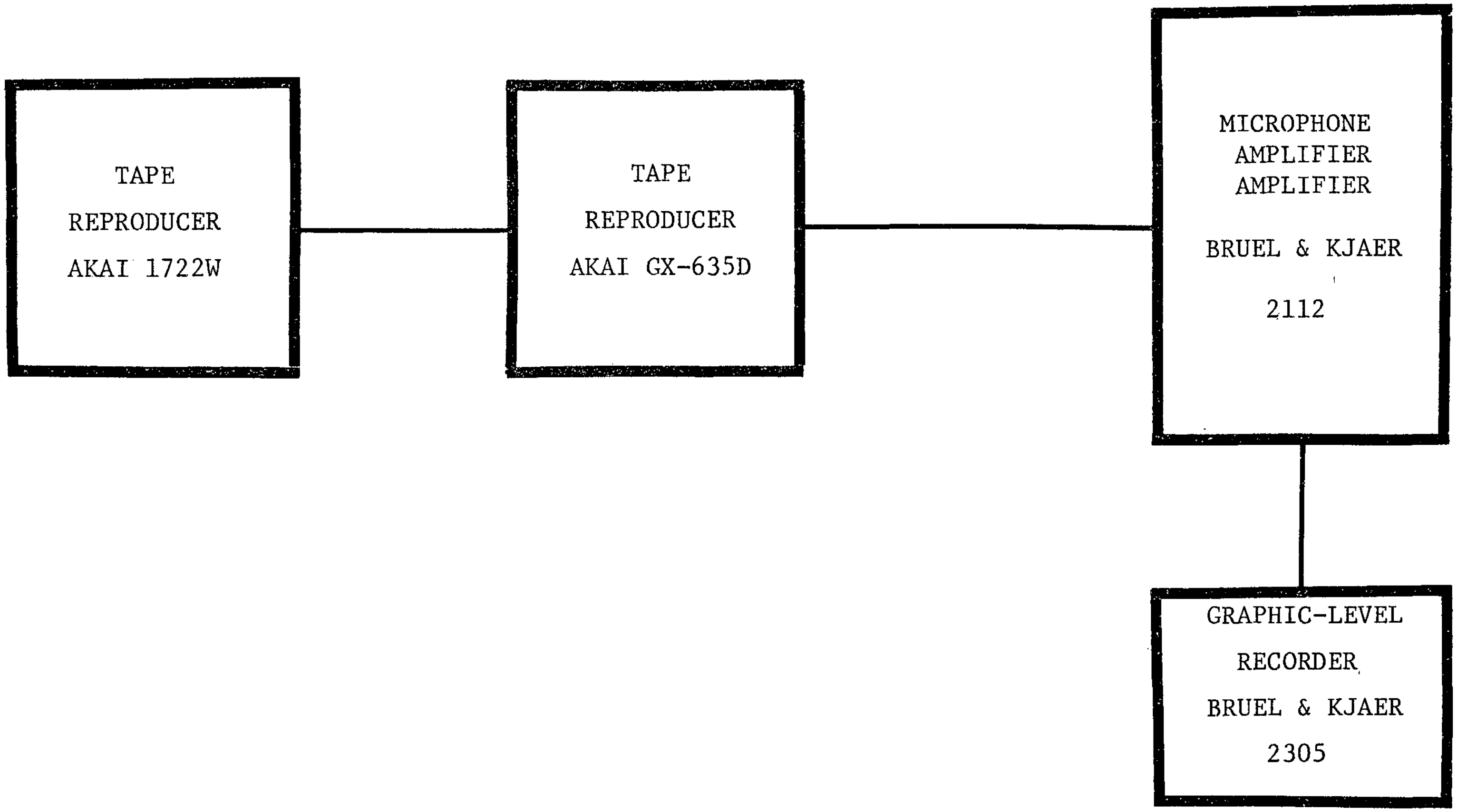

TAPE
REPRODUCER
AKAI 1722W
TAPE
REPRODUCER
AKAI GX-635D
MICROPHONE
AMPLIFIER
AMPLIFIER
BRUEL & KJAER
2112
GRAPHIC-LEVEL
RECORDER
BRUEL & KJAER
2305

FIGURE E1

TABLE E1

Recording Levels

Mean values of recorded level for equal intelligibility as reported by four independent studies (see Table D1) for each of the spondees selected as stimuli for the present investigation. The actual recording level used in the present study is indicated, and was referenced to the mean value of headlight.

SPONDEES	MEAN	RECORDING LEVEL
airplane	5.8	-3.7
armchair	6.5	-3.0
daybreak	5.4	-1.1
headlight	9.5	0.0
horseshoe	9.0	-0.5
inkwell	8.1	-1.7
mousetrap	8.0	-1.5
mushroom	8.9	-0.6
padlock	8.9	-0.6
schoolboy	8.9	-0.6
sidewalk	6.9	-2.6
stairway	8.1	-1.4
sunset	7.9	-1.6
toothbrush	8.7	-0.8
whitewash	8.0	-1.5

FIGURE E2

Hard-copy tracings of the recording level of each stimulus spondee used in the present investigation. Levels recorded on channel 1 of the tape recorder is indicated by the dark tracing, levels recorded on channel 2 of the tape recorder are indicated by the lighter tracing.

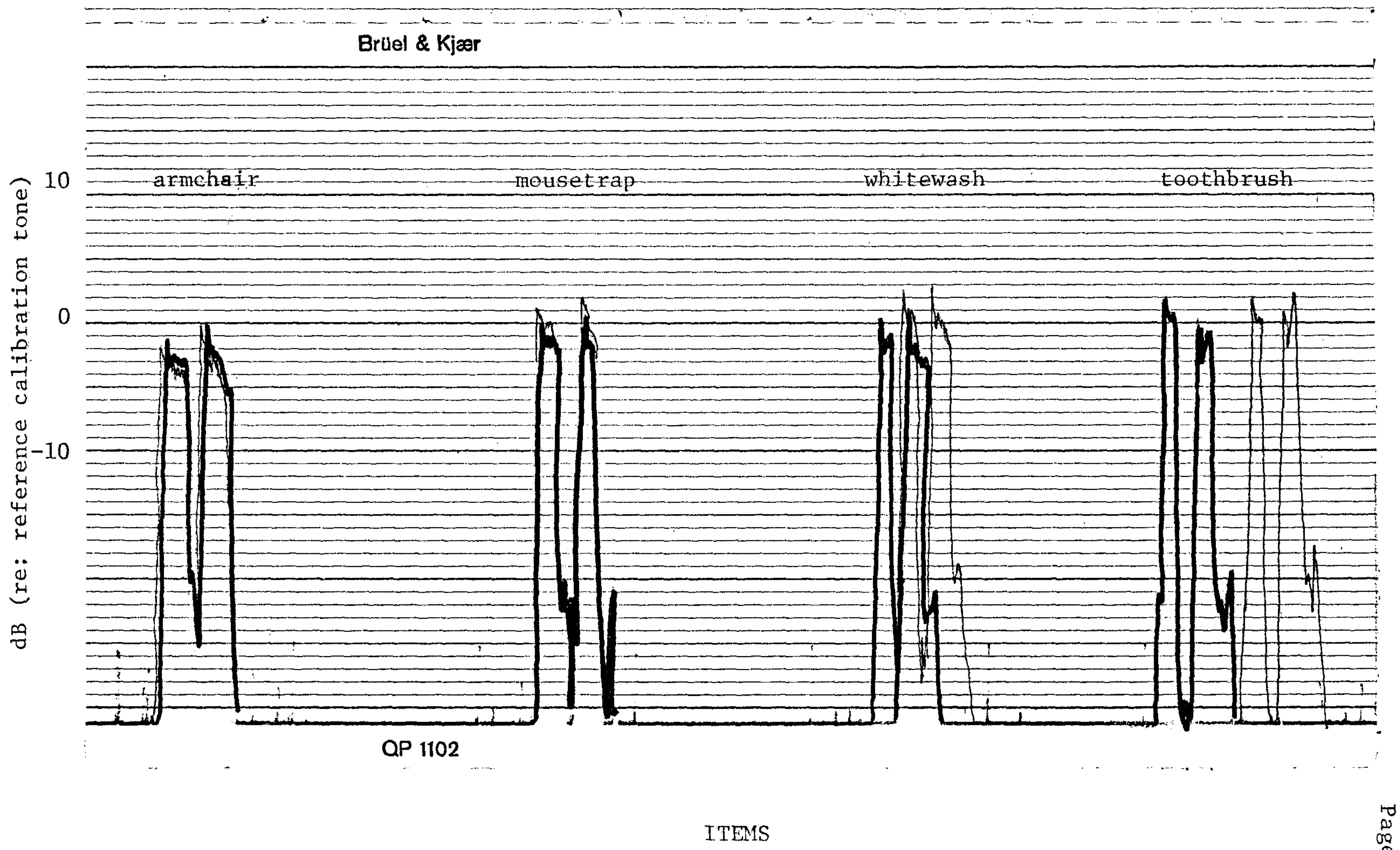

Brüel & Kjær
dB (re: reference calibration tone)
10
0
-10
armchair
mousetrap
whitewash
toothbrush
QP 1102
ITEMS

FIGURE E2

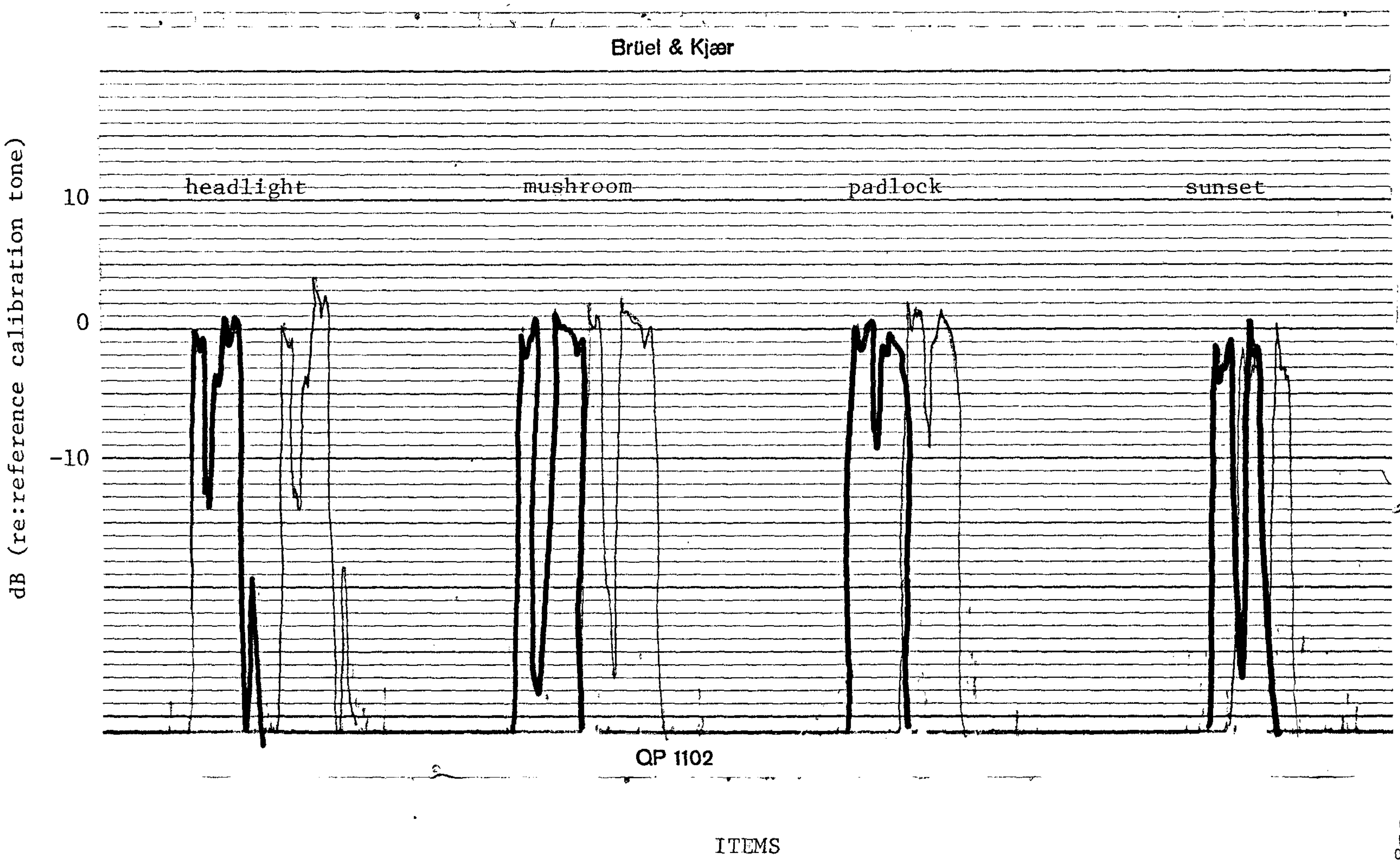


FIGURE E2 (continued)

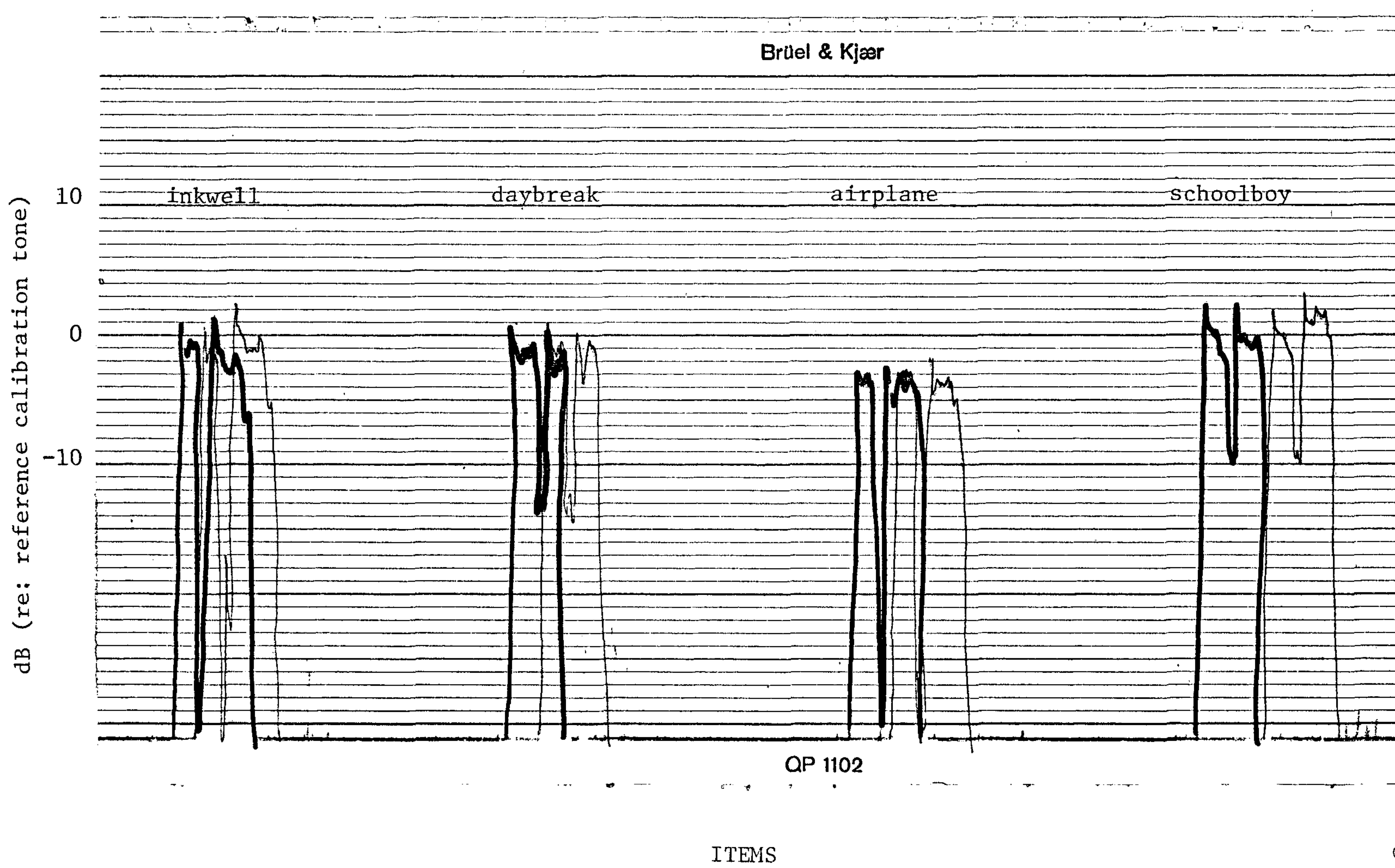


FIGURE E2 (continued)

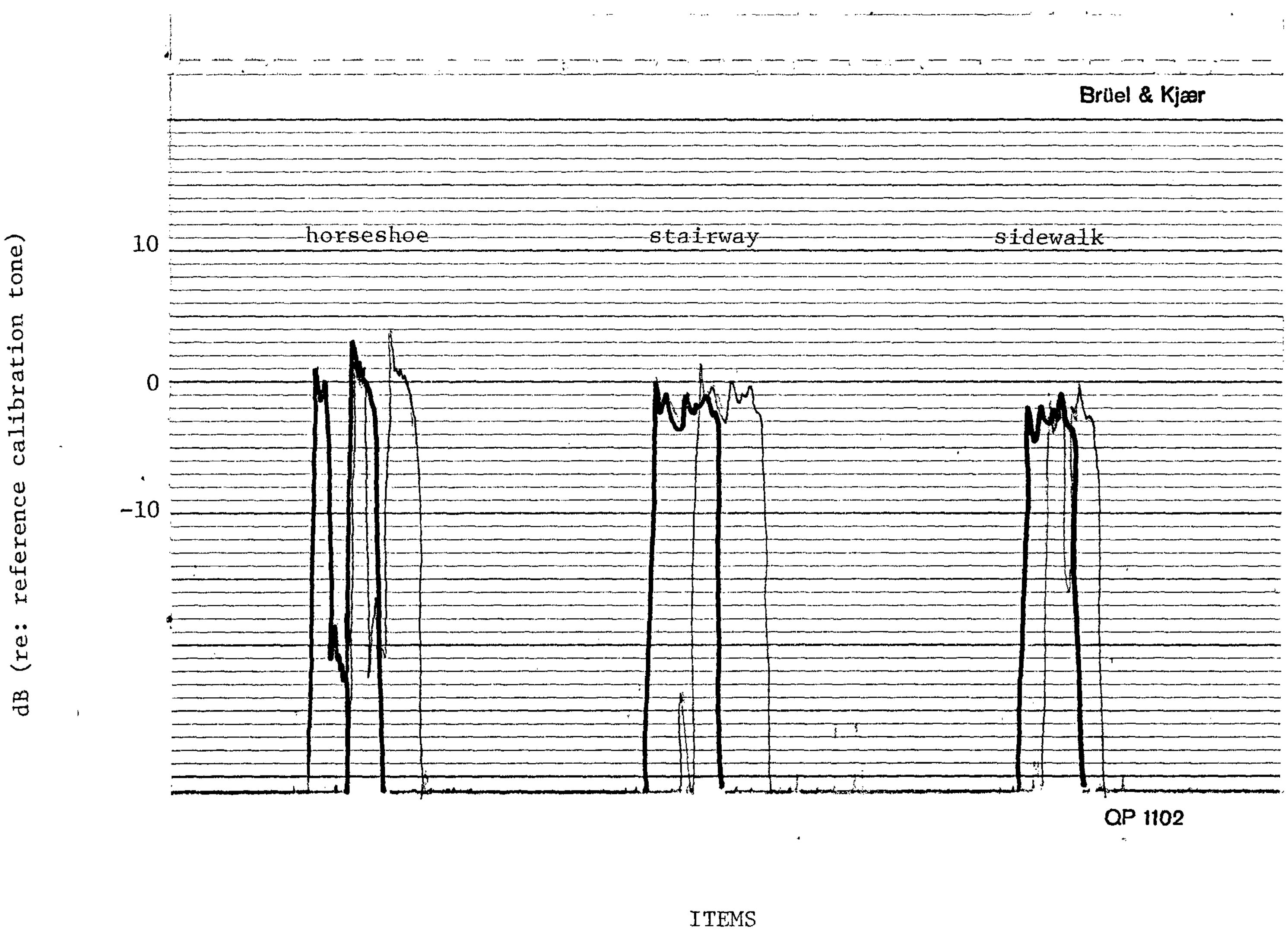


FIGURE E2 (continued)

The master tape containing 15 spondees was then re-recorded to form the five 15-word randomizations that were used in the present investigation. Hard-copy tracings were also made of the second-generation spondees to ensure that their morphology and dB levels conformed to the recording levels of the master tape.

APPENDIX F

INSTRUCTIONS TO THE SUBJECTS

"You will hear a series of two-syllable words. Some will be very soft and difficult to understand, and others will be easy to understand. Repeat the words you hear. If you are unsure, go ahead and guess, or repeat the sounds you hear."

APPENDIX G

PROCEDURE FOR DETERMINATION OF ST

FLOW CHARTS

Figures G1, G2, and G3 contain flow charts summarizing the procedures used in determining ST with ascending, descending and bracketing psychophysical methods with modifications for 2- and 5-dB increment sizes.

FIGURE G1

Flow chart for the ascending psychophysical method.

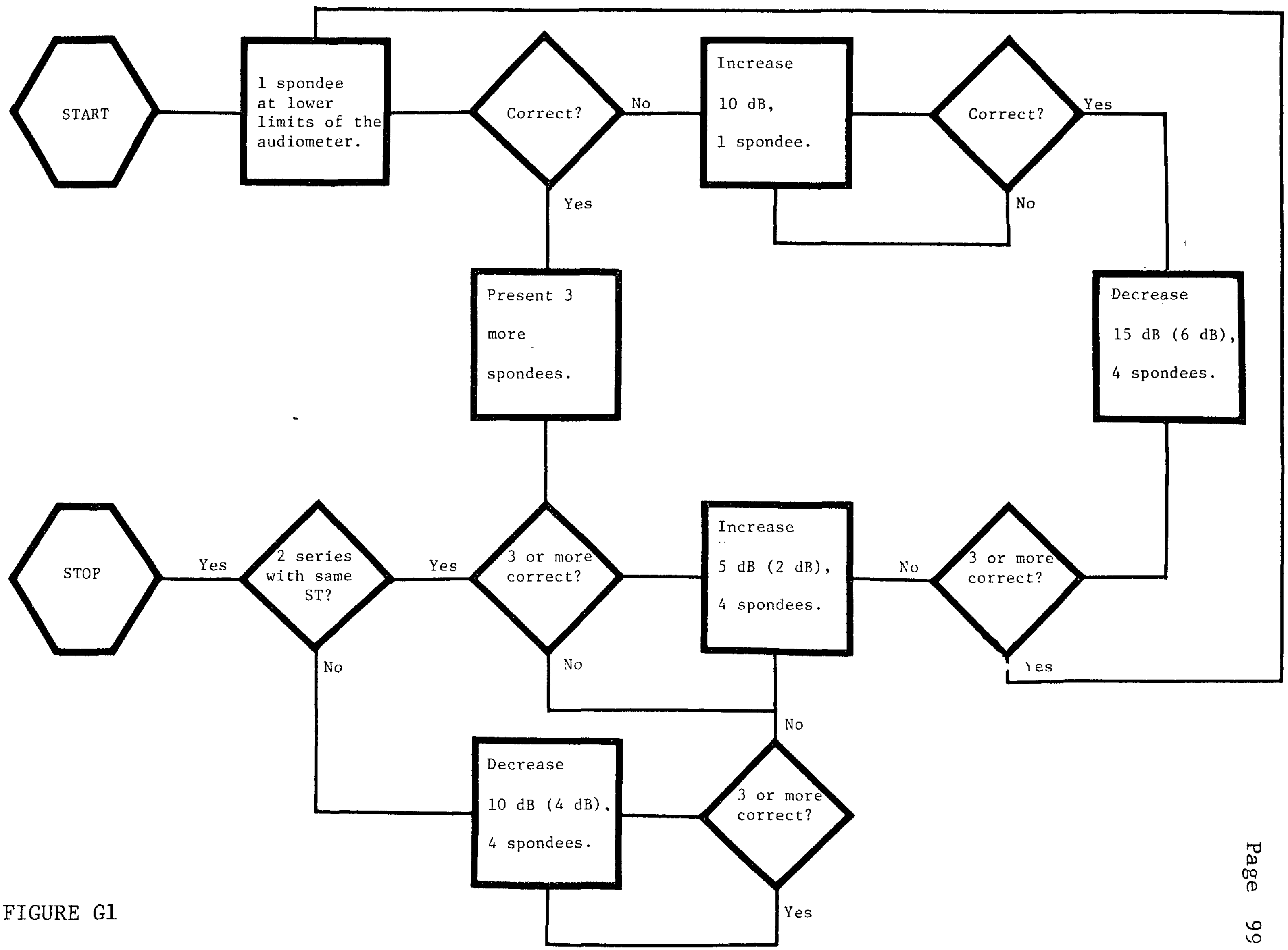

START
1 spondee at lower limits of the audiometer.
Correct?
No
Yes
Increase 10 dB, 1 spondee.
Correct?
Yes
No
Present 3 more spondees.
Decrease 15 dB (6 dB), 4 spondees.
STOP
Yes
2 series with same ST?
Yes
3 or more correct?
Increase 5 dB (2 dB), 4 spondees.
No
3 or more correct?
No
No
Yes
No
Decrease 10 dB (4 dB), 4 spondees.
3 or more correct?
Yes

FIGURE G1

FIGURE G2

Flow chart for the descending psychophysical method.

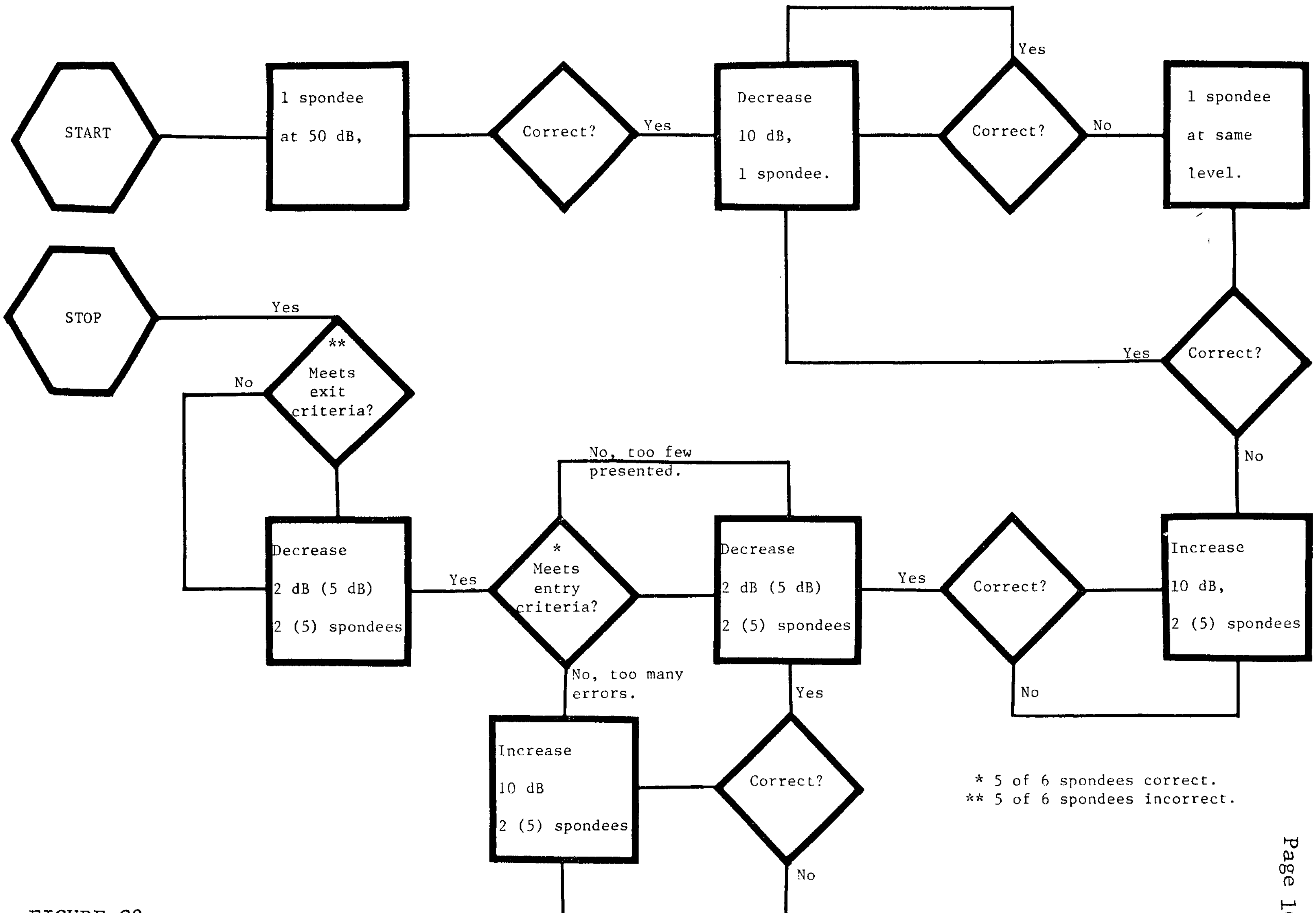
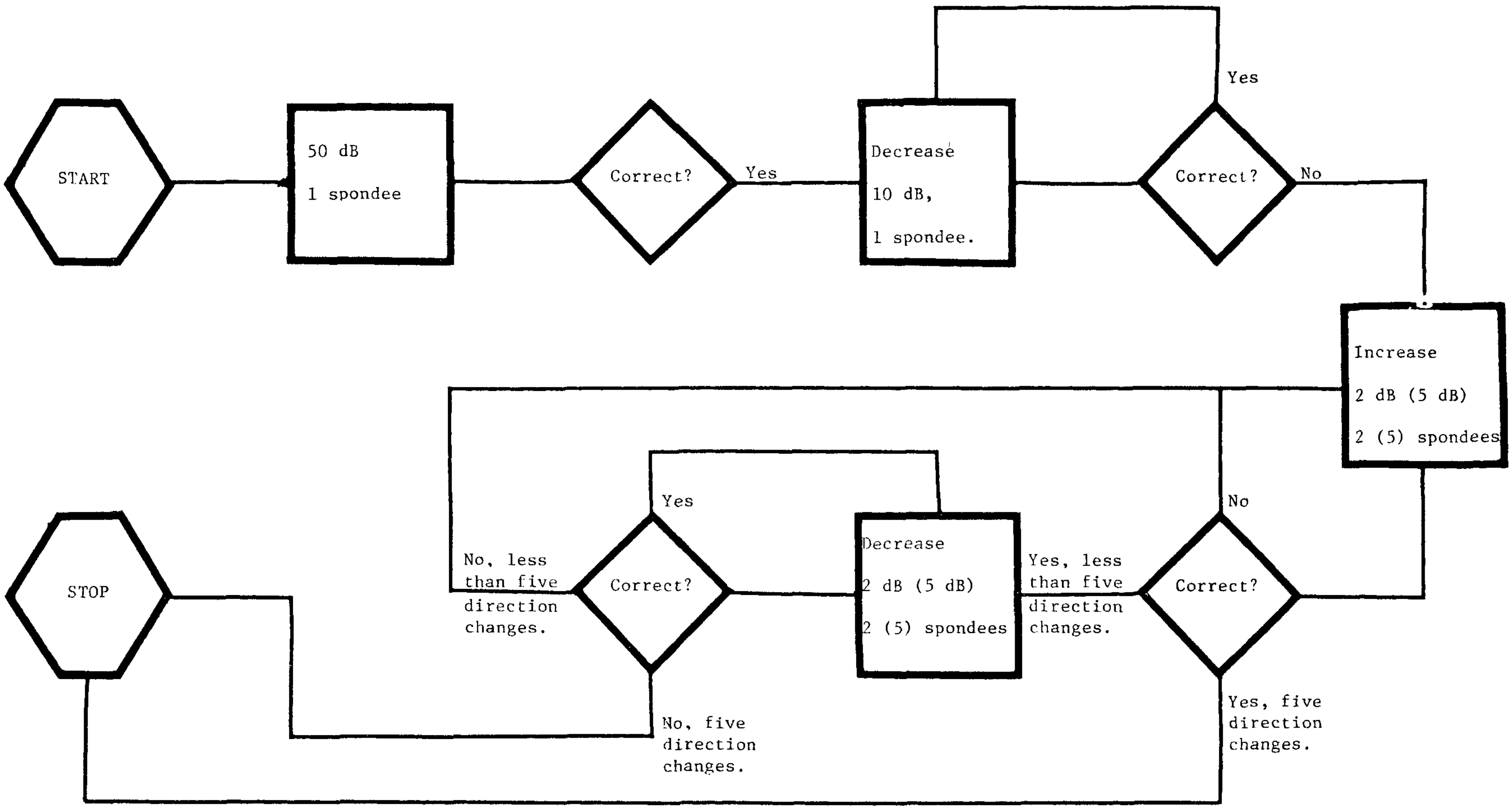

START
1 spondee
at 50 dB,
Correct?
Yes
Decrease
10 dB,
1 spondee.
Yes
Correct?
No
1 spondee
at same
level.
Yes
Correct?
No
STOP
Yes
**
Meets
exit
criteria?
No
No, too few
presented.
Decrease
2 dB (5 dB)
2 (5) spondees
Yes
*
Meets
entry
criteria?
Decrease
2 dB (5 dB)
2 (5) spondees
Yes
Correct?
Increase
10 dB,
2 (5) spondees
No
No, too many
errors.
Yes
Increase
10 dB
2 (5) spondees
Correct?
No
* 5 of 6 spondees correct.
** 5 of 6 spondees incorrect.

FIGURE G2

FIGURE G3

Flow chart for the bracketing psychophysical method.



START
50 dB
1 spondee
Correct?
Yes
Decrease
10 dB,
1 spondee.
Yes
Correct?
No
Increase
2 dB (5 dB)
2 (5) spondees
No
Yes, less
than five
direction
changes.
Correct?
Yes, five
direction
changes.
Decrease
2 dB (5 dB)
2 (5) spondees
Yes
Correct?
No, less
than five
direction
changes.
No, five
direction
changes.
STOP

FIGURE G3

APPENDIX H

COUNTERBALANCING AND RANDOMIZATION

The order in which the 12 ST were obtained for each of the 60 subjects was randomized for increment size and counterbalanced for psychophysical method. Increment size (2- or 5-dB) was randomly assigned for the first three Test Conditions using a table of random numbers. The fourth through sixth Test Conditions were assigned increment values so that within the first six Test Conditions each psychophysical method - increment combination was presented. Test Conditions 7 through 12 followed the same pattern as Test Conditions 1 through 6.

Psychophysical method was counterbalanced so that 10 of the 60 subjects received each possible permutation of the three psychophysical methods. Test Conditions 1 through 6 and 7 through 12 followed the same pattern. A list of the counterbalanced and randomized presentation orders for the 60 subjects is contained in Table H1.

TABLE H1

ORDER OF PRESENTATION

The six orders of presentation are shown with ten subjects being exposed to each of the orders listed.

1. ASCENDING - DESCENDING - BRACKETING
2. DESCENDING - ASCENDING - BRACKETING
3. ASCENDING - BRACKETING - DESCENDING
4. DESCENDING - BRACKETING - ASCENDING
5. BRACKETING - ASCENDING - DESCENDING
6. BRACKETING - DESCENDING - ASCENDING

APPENDIX I

RAW DATA

Data obtained from the 60 subjects are contained in Table I1. Columns 1 through 48 represent the 12 ST for each subject with each ST recorded as a number with 2 digits to the left and one to the right of the decimal point. The ST are recorded so that:

Columns	1 - 4	Method A2T1
	5 - 8	A2T2
	9 - 12	D2T1
	13 - 16	D2T2
	17 - 20	B2T1
	21 - 24	B2T2
	25 - 28	A5T1
	29 - 32	A5T2
	33 - 36	D5T1
	37 - 40	D5T2
	41 - 44	B5T1
	45 - 48	B5T2

(A = ascending; D = descending; B = bracketing; 2 = 2 dB; 5 = 5 dB; T1 = Trial 1; T2 = Trial 2)

Columns 49 and 50 contain a 2 digit subject number. Columns 51 through 55 contain the two-frequency PTA. Columns 56 through 60 contain the three-frequency PTA.

TABLE I1

RAW DATA

```
18.018.015.016.016.015.020.020.016.014.015.015.801005.0005.0
06.004.006.004.006.703.710.010.006.004.003.403.407000.0001.7
04.000.001.0-2.0-2.4-0.700.000.001.0-2.000.0-2.513-02.5-01.7
06.008.010.005.004.004.305.010.003.003.004.205.819005.0005.0
04.002.0-2.0-1.000.000.305.000.009.0-1.004.201.625-02.5-01.6
12.014.010.011.009.710.010.015.009.010.009.209.231007.5008.3
04.004.006.005.001.703.310.010.003.001.003.305.037-02.5000.0
04.002.000.004.000.300.705.005.003.000.001.7-0.843-02.5000.0
10.008.008.008.007.007.710.010.010.007.005.806.749005.0006.7
04.002.002.000.0-1.3-2.705.005.0-1.0-1.000.001.755-10.0-06.7
04.002.000.002.002.702.010.005.001.0-1.0-0.801.702-02.5000.0
04.004.001.005.003.403.705.005.000.002.005.004.208002.5003.4
00.002.0-1.001.001.7-1.705.005.000.0-1.0-0.800.014-05.0-03.4
12.012.007.012.010.308.715.015.013.008.014.210.020012.5015.0
06.010.004.005.004.004.310.010.003.004.005.005.026002.5003.3
02.002.002.001.0-2.300.605.000.001.0-1.003.300.032-02.5-01.3
08.008.011.005.006.707.715.010.004.008.007.506.738002.5005.0
02.002.003.0-3.0-1.7-1.705.000.001.0-2.000.8-1.744-07.5-05.0
00.000.001.0-2.001.000.010.000.001.0-2.000.0-0.850000.0001.7
06.004.004.003.005.703.305.005.008.005.005.805.056002.5003.3
04.002.002.001.002.702.705.005.001.001.004.204.203000.0001.7
04.008.003.002.004.003.010.005.002.001.006.702.509000.0001.7
04.000.003.0-1.003.002.005.005.003.006.001.7-0.815-02.5-02.5
08.004.002.004.004.003.010.010.004.005.005.005.021000.0001.7
16.016.015.014.014.016.320.020.013.015.015.014.227010.0010.0
04.002.002.003.004.000.710.005.002.003.002.501.733002.5008.3
04.002.003.001.002.7-1.005.005.002.000.001.7-2.539000.0000.0
02.004.003.004.004.700.600.000.000.0-2.002.500.845000.0000.0
06.008.005.006.006.304.315.010.006.004.005.805.851000.0000.0
00.000.0-1.0-2.0-1.3-2.305.005.0-3.0-2.000.8-2.557-05.0-05.0
14.008.006.007.006.707.410.010.005.006.003.304.204005.0006.7
10.008.004.008.006.005.710.005.007.004.007.506.710-02.5000.0
04.006.004.005.002.003.010.010.002.004.003.401.716-02.5003.4
12.010.011.006.007.008.310.010.014.009.009.208.322005.0008.3
02.002.005.002.005.303.010.005.003.001.003.303.328002.5003.3
08.008.006.006.004.004.315.010.007.008.004.205.034000.0008.3
04.004.003.001.004.001.005.005.003.002.000.801.740007.5008.3
00.004.003.003.0-2.701.010.005.0-3.000.000.000.046-05.0001.7
02.002.000.001.000.600.005.005.001.0-5.000.000.052000.0001.7
04.004.006.004.000.301.010.005.0-1.0-1.000.000.858002.5003.3
```

RAW DATA (continued)

```
06.004.008.005.003.703.010.005.001.002.006.704.205005.0005.0
08.008.010.010.008.408.410.015.005.008.006.708.411-02.5003.4
12.010.007.008.008.705.310.010.008.006.005.805.017-02.5001.7
08.008.007.002.002.304.010.005.004.006.005.003.323005.0006.3
06.004.005.002.004.000.310.005.004.003.002.501.729002.5003.3
06.006.003.005.003.705.310.010.003.003.002.500.835000.0001.7
04.000.003.0-1.002.0-0.705.000.004.004.000.800.841007.5008.3
02.000.004.0-1.000.3-2.605.000.0-3.0-5.0-1.7-2.547-02.5-01.7
04.006.005.002.001.005.310.005.001.002.001.703.353000.0001.7
02.004.004.000.003.3-1.610.000.002.002.000.0-1.659002.5005.0
08.006.005.005.004.304.010.005.005.003.005.006.706005.0006.7
08.004.006.005.004.704.410.010.007.003.009.205.912010.0013.4
06.004.003.002.001.702.710.010.003.004.005.000.018-02.5-01.7
08.006.005.005.004.704.005.010.006.003.006.705.824-02.5000.0
08.008.004.002.006.302.010.005.004.001.005.003.330000.0001.6
06.006.004.003.0-0.300.005.005.000.002.000.800.836-02.5000.0
-2.002.0-2.002.0-4.7-0.700.005.0-2.000.0-3.300.842000.0000.0
12.008.009.005.011.306.010.010.007.003.009.206.748012.5013.3
06.004.002.001.005.000.005.005.003.003.001.702.554002.5003.3
06.004.002.0-3.001.3-2.005.005.000.0-1.001.701.760-05.0-01.6
```

APPENDIX J

ANOVA

Analyses of variance (ANOVA) were obtained for ST, ST minus the two-frequency PTA and ST minus the three-frequency PTA using a computer program developed by Ullrich and Pitz (1981). A computer (DECsystem-20) was used in this analysis. Tables J1 through J8 contain a summary of ANOVA results.

TABLE J1

SPONDEE THRESHOLD ANOVA SUMMARY TABLE

Order (O [see Table 22]) by Increment Size (I [2 dB versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) by Trial (T [trial 1 versus trial 2]) analyses of variance summary table for spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
O	210.421	42.0841	5	0.259	0.9326
Error	8785.260	162.690	54		
OxI	16.9270	3.3854	5	0.997	0.5708
Error	183.392	3.3962	54		
OxP	44.3424	4.4342	10	1.444	0.1705
Error	331.650	3.0708	108		
OxIxP	30.8813	3.0881	10	0.723	0.7022
Error	461.014	2.2685	108		
OxT	10.6630	2.1326	5	0.370	0.8675
Error	311.436	5.7673	54		
IxT	5.0166	5.0166	1	1.776	0.1852
Error	152.511	2.8243	54		
OxIxT	22.9874	4.5974	5	1.628	0.1676
Error	152.511	2.8243	54		
PxT	6.4912	3.2456	2	1.805	0.1674
Error	194.242	1.7985	108		
OxPxT	41.0449	4.1045	10	2.282	0.0180
Error	194.242	1.7985	108		

TABLE J1 (continued)

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
IxPxT	9.9908	4.9954	2	1.612	0.2025
Error	334.726	3.0993	108		
OxIxPxT	60.6173	6.0617	10	1.956	0.0449
Error	334.726	3.0993	108		

TABLE J2

BRACKETING METHOD ANOVA SUMMARY TABLE
FOR SPONDEE THRESHOLD

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the bracketing psychophysical method for spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	1.4520	1.4520	1	0.592	0.5489
Error	144.758	2.4535	59		

TABLE J3

DESCENDING METHOD ANOVA SUMMARY TABLE FOR SPONDEE THRESHOLD

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the descending psychophysical method for spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
I	17.6333	17.6333	1	3.475	0.0693
Error	299.367	5.0740	59		

TABLE J4

ANOVA SUMMARY TABLE FOR THE NUMBER OF
OF SPONDEES REQUIRED FOR SPONDEE THRESHOLD

Order (O [see Table 22]) by Increment Size (I [2 dB versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) by Trial (T [trial 1 versus trial 2]) analyses of variance for the number of spondees required to obtain a spondee threshold.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
O	658.550	131.710	5	0.706	0.6236
Error	10071.80	186.515	54		
OxI	787.794	157.559	5	1.731	0.1426
Error	4915.76	91.0327	54		
OxP	1529.19	152.919	10	1.596	0.1171
Error	10349.1	95.8247	108		
OxIxP	716.618	71.6618	10	0.765	0.6633
Error	10117.9	93.6846	108		
T	231.200	231.200	1	3.050	0.0827
Error	4093.07	75.7975	54		
OxT	576.900	115.380	5	1.522	0.1975
Error	4093.07	75.7975	54		
IxT	9.7993	9.7993	1	0.121	0.7289
Error	4359.54	80.7322	54		
OxIxT	326.833	65.3666	5	0.810	0.5494
Error	4399.54	80.7322	54		
PxT	100.208	50.1038	2	0.767	0.5204
Error	705.365	65.3116	108		

TABLE J4 (continued)

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
OxPxT	814.494	81.4494	10	1.247	0.2693
Error	7053.65	65.3116	108		
IxPxT	112.006	56.0029	2	0.839	0.5617
Error	7209.36	66.7533	108		
OxIxPxT	781.958	78.1958	10	1.171	0.3175
Error	7209.36	66.7533	108		

TABLE J5

SPONDEE THRESHOLD/TWO FREQUENCY AGREEMENT
ANOVA SUMMARY TABLE

Order (O [see Table 22]) by Increment Size (I [2 dB versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) by Trial (T [trial 1 versus trial 2]) analyses of variance summary table for the agreement between spondee threshold and the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
O	705.554	141.111	5	0.880	0.5021
Error	8657.42	160.323	54		
OxT	18.3761	3.6752	5	1.118	0.3618
Error	177.488	3.2868	54		
OxP	47.2063	4.7206	10	1.430	0.1765
Error	356.604	3.3019	108		
OxIxP	36.8353	3.6835	10	0.854	0.5785
Error	465.588	4.3110	108		
OxT	6.6174	1.3234	5	0.236	0.9438
Error	303.078	5.8166	54		
IxT	6.0317	6.0317	1	2.114	0.1481
Error	154.086	2.8534	54		
OxIxT	26.1610	5.2322	5	1.834	0.1211
Error	154.086	2.8535	54		
PxT	4.1709	2.0854	2	1.161	0.3172
Error	194.036	2.9679	108		
/	/	/	/	/	/

TABLE J5 (continued)

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
OxPxT	41.0350	4.1035	10	2.284	0.0179
Error	194.036	2.9679	108		
IxPxT	6.7017	3.3509	2	1.129	0.3275
Error	320.529	2.9679	108		
OxIxPxT	49.4187	4.9419	10	1.665	0.0979
Error	320.529	2.9679	108		

TABLE J6

BRACKETING METHOD ANOVA SUMMARY TABLE
FOR SPONDEE THRESHOLD/TWO FREQUENCY AGREEMENT

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the agreeement between spondee threshold and the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	mean SQUARE	DF	F-RATIO	PROB.
I	0.2521	0.2521	1	0.106	0.7445
Error	140.023	2.3732	59		

TABLE J7

DESCENDING METHOD ANOVA SUMMARY TABLE FOR SPONDEE THRESHOLD/TWO FREQUENCY AGREEMENT

Simple-effects analysis of variance for Increment Size (I [2 dB versus 5 dB]) for the agreeement between spondee threshold and the two-frequency pure-tone average.

SOURCE	SUM OF SQUARES	mean SQUARE	DF	F-RATIO	PROB.
I	17.7870	17.7870	1	3.734	0.0550
Error	281.033	4.7633	59		

TABLE J8

SPONDEE THRESHOLD/THREE FREQUENCY AGREEMENT
ANOVA SUMMARY TABLE

Order (O [see Table 22]) by Increment Size (I [2 dB versus 5 dB]) by Psychophysical Method (P [ascending versus descending versus bracketing]) by Trial (T [trial 1 versus trial 2]) analyses of variance summary table for the agreement between spondee threshold and the three-frequency pure-tone average.

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
O	1306.64	261.328	5	1.923	0.1049
Error	7338.99	125.907	54		
OxI	28.1984	5.6397	5	1.546	0.1905
Error	197.018	3.6485	54		
OxP	48.5313	4.8531	10	1.379	0.1989
Error	379.964	3.5181	108		
OxIxP	29.0528	2.9053	10	0.997	0.7264
Error	449.987	4.1665	108		
OxT	16.9672	3.3935	5	0.620	0.6874
Error	295.453	5.4714	54		
IxT	6.1791	6.1791	1	2.241	0.1368
Error	148.894	2.7573	54		
OxIxT	20.9657	4.1532	5	1.506	0.2025
Error	148.894	2.7573	54		
PxT	10.8539	5.4269	2	2.766	0.0656
Error	211.923	1.9623	108		
IxPxT	15.2399	7.6188	2	2.352	0.0980
Error	349.913	3.2399	108		

APPENDIX K

CORRELATIONS

Correlation coefficients were obtained for four sets of variables: ST versus two- and three-frequency PTA for Trial 1, ST versus two-and three-frequency PTA for Trial 2; ST for Trial 1 versus Trial 2; and two-frequency PTA versus three-frequency PTA. Correlation matrices are contained in Tables K1 through K4.

TABLE K1

CORRELATION COEFFICIENTS FOR PURE-TONE AVERAGES AND TEST CONDITIONS

Pearson Product-Moment correlation coefficients are shown for two- and three-frequency pure-tone averages with each test condition (ascending [A], descending [D] and bracketing [B] by 2-dB [2] versus 5-dB [5] increment sizes) for Trial 1.

	2 FPTA	3 FPTA
A2	0.5484*	0.5718*
D2	0.5569*	0.5830*
B2	o.6292*	0.6198*
A5	0.4128**	0.5096*
D5	0.5664*	0.5830*
B5	0.5896*	0.5942*

* $p < 0.000$
** $p < 0.001$

TABLE K2

CORRELATION COEFFICIENTS FOR PURE-TONE AVERAGES AND TEST CONDITIONS

Pearson Product-Moment correlation coefficients are shown for two- and three-frequency pure-tone averages with each test condition (ascending [A], descending [D] and bracketing [B] by 2-dB [2] versus 5-dB [5] increment sizes) for Trial 2.

	2 FPTA	3 FPTA
A2	0.4916*	0.5262*
D2	0.5044*	0.5587*
B2	o.5682*	0.5866*
A5	0.4128**	0.4715*
D5	0.5358*	0.5889*
B5	0.5673*	0.5948*

* $p < 0.000$
** $p < 0.001$

TABLE K3

CORRELATION COEFFICIENTS FOR REPEATED MEASURES

Pearson Product-Moment correlation coefficients are shown between trial 1 and trial 2 for each test condition (ascending [A], descending [D] and bracketing [B] by 2-dB [2] versus 5-dB [5] increment sizes).

METHOD	*r*
A2 (T1 vs T2)	0.8447
D2 (t1 vs T2)	0.7545
B2 (t1 vs T2)	0.8357
A5 (t1 vs T2)	0.6967
D5 (t1 vs T2)	0.8045
B5 (t1 vs T2)	0.8861

TABLE K4

CORRELATION COEFFICIENTS FOR PURE-TONE AVERAGES

Pearson Product-Moment correlation coefficients are shown for the two- (2 FPTA) with the three-frequency (3 FPTA) pure-tone averages.

	2 FPTA	3 FPTA
2 FPTA	1.000*	0.924*
3 FPTA	0.924*	1.000*

* $p < 0.000$

APPENDIX L

TESTS FOR DIFFERENCES AMONG CORRELATION COEFFICIENTS

A t-test for the difference between dependent correlation coefficients (Brunig and Kintz, 1978) was used to determine whether correlations contained in Appendix K were different at the 0.01 level of confidence. Table L1 contains the results of these t-tests.

TABLE L1

T-TEST FOR DIFFERENCE BETWEEN DEPENDENT CORRELATIONS

Significance of differences between correlation coefficients are shown for the repeated measures as a function of test condition (ascending [A], descending [D] and bracketing [B] by 2-dB [2] versus 5-dB [5] increments).

	TRIAL 1	TRIAL 2
A2	0.56362	0.78908
D2	0.55837	1.27042
B2	0.23648	0.44211
A5	2.21652	1.29370
D5	0.39775	1.27445
B5	0.11187	0.66443

$t > 2.390$ is significant at 0.01 level

APPENDIX M

RANGES: ST MINUS THE TWO-FREQUENCY PTA

Table M1 contains a comparison of the range of agreement for ST minus the two-frequency PTA.

TABLE M1

RANGES of ST

Percentages of sample yielding spondee thresholds/two-frequency pure-tone average agreements within specified ranges. Negative numbers in the ranges indicate a more sensitive spondee threshold than would be expected from the pure-tone average. Methods are identified by psychophysical method (ascending [A], descending [D], or bracketing [B]), Increment Size (2 dB [2] or 5 dB [5]) and Trial (trial 1 [T1], or trial 2 [T2]).

METHOD	<-10	-5 to -10	±5	+5 to +10	>+10
A2T1	0.0%	0.0%	63.4%	26.7%	10.0%
A2T2	0.0%	3.4%	63.4%	26.7%	6.7%
D2T1	0.0%	1.7%	75.0%	18.4%	5.0%
D2T2	0.0%	5.0%	78.4%	10.0%	6.7%
B2T1	0.0%	3.4%	75.0%	16.7%	5.0%
B2T2	0.0%	6.7%	73.4%	16.7%	3.4%
A5T1	0.0%	0.0%	40.0%	45.0%	15.0%
A5T2	0.0%	1.7%	58.4%	30.0%	10.0%
D5T1	0.0%	1.7%	73.4%	21.7%	3.4%
D5T2	0.0%	5.0%	81.7%	11.7%	1.7%
B5T1	0.0%	3.4%	75.0%	20.0%	0.0%
B5T2	0.0%	5.0%	80.0%	11.7%	3.4%

APPENDIX N

RANGES: ST MINUS THE THREE-FREQUENCY PTA

Table N1 contains a comparison of the range of agreement for ST minus the three-frequency PTA.

TABLE N1

RANGES of ST

Percentages of sample yielding spondee thresholds/three-frequency pure-tone average agreements within specified ranges. Negative numbers in the ranges indicate a more sensitive spondee threshold than would be expected from the pure-tone average. Methods are identified by psychophysical method (ascending [A], descending [D], or bracketing [B]), Increment Size (2 dB [2] or 5 dB [5]) and Trial (trail 1 [T1], or trial 2 [T2]).

METHOD	<-10	-5 to -10	±5	+5 to +10	>+10
A2T1	0.0%	1.7%	70.0%	23.4%	5.0%
A2T2	0.0%	6.7%	68.4%	23.4%	1.7%
D2T1	0.0%	8.4%	76.7%	15.0%	0.0%
D2T2	0.0%	8.4%	81.7%	8.4%	1.7%
B2T1	0.0%	3.4%	85.0%	10.0%	1.7%
B2T2	0.0%	11.7%	83.4%	5.0%	0.0%
A5T1	0.0%	0.0%	41.7%	51.7%	6.7%
A5T2	0.0%	1.7%	66.7%	28.4%	3.4%
D5T1	0.0%	6.7%	76.7%	13.4%	3.4%
D5T2	0.0%	8.4%	86.7%	5.0%	0.0%
B5T1	0.0%	5.0%	81.7%	13.4%	0.0%
B5T2	0.0%	8.4%	81.7%	6.7%	1.7%

APPENDIX O

RANGES: TRIAL 1 VERSUS TRIAL 2

Table O1 contains a comparison of the ranges of agreement between Trial 1 and Trial 2 for ST minus the two-frequency PTA and ST minus the three-frequency PTA.

TABLE 01

RANGES of ST

Percentages of sample yielding spondee threshold trial 1 versus trial 2 agreements within specified ranges. Negative numbers in the ranges indicate a more sensitive spondee threshold for trial 1 than for Trial 2. Methods are identified by psychophysical method (ascending [A], descending [D], or bracketing [B]), and Increment Size (2 dB [2] or 5 dB [5]).

METHOD	<-10	-5 to -10	±5	+5 to +10	>+10
A2	0.0%	0.0%	98.3%	1.7%	0.0%
D2	0.0%	0.0%	96.7%	3.3%	0.0%
B2	0.0%	0.0%	100.0%	0.0%	0.0%
A5	0.0%	0.0%	100.0%	0.0%	0.0%
D5	0.0%	0.0%	93.3%	0.0%	1.7%
B5	0.0%	0.0%	100.0%	0.0%	0.0%