1981

Accuracy of recorded measures of the Willeford Central Auditory Processing test battery

Susan L. Shea
The University of Montana

ACCURACY OF RECORDED MEASURES

OF THE WILLEFORD CENTRAL

AUDITORY PROCESSING TEST BATTERY

by

Susan L. Shea

B.Sc., Speech Pathology & Audiology
University of Alberta, 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

June, 1981

Approved by:

Michael J. McRaffin

Chairman, Board of Examiners

Dean, Graduate School

7 8 81

Date

ACCURACY OF RECORDED MEASURES

OF THE WILLEFORD CENTRAL

AUDITORY PROCESSING TEST BATTERY

by

Susan L. Shea

B.Sc., Speech Pathology & Audiology
University of Alberta, 1979

An Abstract

Of 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

June, 1981

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

ABSTRACT

Shea, Susan L., M.A., December, 1981 Communication Sciences & Disorders

Accuracy of recorded measures of the Willeford Central Auditory Processing Test Battery (125 pp.)

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

Thesis approved: ____________________

A frequently used test for the assessment of central auditory processing in children is the Willeford Central Auditory Processing (CAP) test battery. On the basis of results from this test, children have been labelled learning disabled, and/or placed in special education curricula. An evaluation of eight magnetic tape recordings of the CAP battery was completed. Findings indicate significant inter-tape differences, inter-test differences, and inter-channel differences in recording levels of test items. All levels were referenced to the calibration tone at the beginning of each tape. The Low-Pass Filtered Speech (LPFS), the Competing Sentences (CST) and the Rapidly Alternating Speech Perception (RASP) tests typically were within 10 dB of the reference tone with the exception of certain items. Inter-channel differences of 38.4 dB were not uncommon on the Binaural Fusion (BF) test, with the high-pass channel often being within the noise floor of the tape.

The findings of the present study indicate that the quality control of the magnetic tape recordings of Willeford's CAP test battery is inconsistent. It is therefore likely that, based strictly on the recording levels of some items, patients as well as normal listeners will perform poorly on the tests involved. In addition, there is little consistency between various tape recordings of the test, thus precluding the rational usage of this test clinically, when a comparison of results between facilities (or any comparison involving the comparison of scores obtained on different tapes) is desired.

ACKNOWLEDGMENTS

The development and completion of this study were accomplished with the help of many individuals. I wish to thank the thesis committee for their suggestions and constructive criticisms in the development of the study; those individuals who kindly provided the tapes used in this study; and family and friends for their encouragement and support. In particular, I would like to extend my appreciation to Dr. Michael J. M. Raffin, who directed this research, for his encouragement, guidance and infinite patience in seeing this study to its completion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

INTRODUCTION

The Central Auditory Processing (CAP) test battery (Willeford, 1976) was developed as a diagnostic tool for the assessment of central-auditory pathology in young children with normal peripheral auditory function. This test battery since has been used, in whole or in part, to evaluate the integrity of the central auditory nervous system (CANS) in persons presenting other primary problems such as peripheral hearing loss, temporal-lobe tumors, decompression sickness, learning disabilities, dyslexia and stuttering (Lynn and Gilroy, 1972; Lynn et al., 1972; Dempsey, 1977; Pinheiro, 1977; Schultz, 1977; Mittenberger, et al., 1978, 1979; Musiek and Geurkink, 1980; Welsh, et al., 1980; Larimore and Willeford, 1980). Subjects ranged in age from 3 years (Logan, 1979) to 65 years (Mittenberger, et al., 1978). Appendix A details the underlying mechanisms associated with this test.

A review of published normative data for both children (White, 1977; Willeford, 1978) and adults (Ivey, 1969; Mittenberger, et al., 1979) reveals some discrepancies between norms reported by Willeford (1978) and those reported by other researchers. A detailed review of pertinent research, and of the test and associated administrative procedures, may be found in Appendix A.

Adequate and justifiable reasons for the observed discrepancies in the foregoing research have not been developed or iterated in the published literature. An indication of a possible contributing factor was observed when clinicians at the University of Montana reported that VU-meter deflections for many items were inconsistent with that caused by the reference calibration pure tone of the tape. For example while some items caused maximum deflection of the VU-meter indicator (after the reference calibration tone had been adjusted to produce a meter deflection of 0), other items caused no discernible meter-indicator movement, thus leading one to conclude that inter-item intensity differences might be in the order of 23 dB (i.e.: the power of some items was nearly 200 fold the power for other items). However, since the clinic tape was several years old, a new tape was purchased directly from Willeford, and the electromagnetic characteristics of that tape were evaluated after it was found that it produced similar VU-meter behavior. These meticulous observations might lead the objective observer to conclude that there may exist some significant discrepancies in the recording levels used by Willeford on the master tape. If this is the case, then test interpretation becomes extremely difficult especially when one attempts to compare clinical results with those published as normative data produced with other tapes, unless the tape-recording levels are identical which they are not.

Thus, the purpose of the present study is to determine whether differences in recording levels of the Willeford CAP test battery exist between items (within tapes), between tests (within tapes), or amongst tests (between tapes).

METHOD

Test Materials: Eight alleged original tape recordings of the CAP test battery were obtained from various clinics, selected randomly and representing a geographical cross-section of the country, and were accepted for inclusion in the present study only if they were less than 3 years old. Two of these tapes had never been used. Details of the tape selection and categorization procedures may be found in Appendix B.

Apparatus and Calibration: A two- or four-track tape recorder was used as the playback instrument for the recordings depending upon the characteristics of the tape recording employed. The output of the tape recorder was fed to an impedance-matched loudspeaker, and output levels were monitored across the loudspeaker with a microphone amplifier (Bruel & Kjaer, Type 2603). The output of the microphone amplifier was connected to the input of a continuously recording graphic-level recorder (Bruel & Kjaer, Type 2305). The tape recorder was cleaned and degaussed prior to any measurement, and also was calibrated to conform to specifications promulgated by the National Association of Broadcasters (NAB, Standard Magnetic Tape Recording and Reproducing, 1965) prior to any measurement of each tape. Details of the calibration may be found in Appendix C. Appendix D contains samples of the hard-copy tracings obtained on one tape.

Procedure: Both channels of each tape were set to a convenient reference point of 0 dB relative to the reference calibration pure tone provided on that tape, prior to obtaining measurements for each of the tests. Since the RASP is preceded by a different calibration pure tone, the reference point of 0 dB was readjusted relative to this calibration tone. Measurements were recorded for each channel, where appropriate, and a permanent tracing made of the amplitude of each recording for each of the four tests. A score was computed for each component of each test as follows: CST -- an average of the total rms amplitude in each sentence; LPFS -- the rms amplitude for each word; BF -- the rms amplitude for each syllable of each spondaically stressed word; RASP -- an average of the total instantaneous peak amplitudes (not rms) for the phonemes in each sentence.

RESULTS

Analyses of variance (ANOVA) performed for each of the four tests - CST, LPFS, BF, and RASP - revealed significant differences in recorded levels between items within tests, between tests within tapes, and between tapes within tests. The raw data upon which these analyses are based are contained in Appendix E. Details of the analyses may be found in Appendix F. For the purposes of the present study, the level of significance was rounded to the next greater whole number to allow the reader to decide whether this level is great enough for his/her needs. Significant inter-channel differences also were noted on the BF test. While differences were significant beyond the 0.01 level of significance, mean item intensity differences were within 10.0 dB for all but the BF test; for the latter, a difference in mean intensity levels between Channel 1 and Channel 2 was in the order of 16.6 dB. Ranges of measured intensity levels for individual items within and between tapes revealed a greater disparity in recorded intensity levels. In addition, one tape (Tape #1) was measured on three separate occasions by one examiner to provide an estimate of intra-examiner reliability. The largest discrepancy for any item of any test between these three measurements was 3 dB.. Findings for each of the four tests are as follows:

CST

Figure 1 depicts a significant three-way interaction that occurred between items, tests and channels. Channel 1 revealed a greater mean intensity range (-1.9 to 4.0; difference of 5.9 dB) between items than Channel 2 (-0.62 to 3.6; difference of 4.22 dB) by approximately 1.7 dB. The largest difference between items occurred for Test 2, Channel 1 and was 5.4 dB.

Figure 1

CST: Test by Item by Channel interaction.

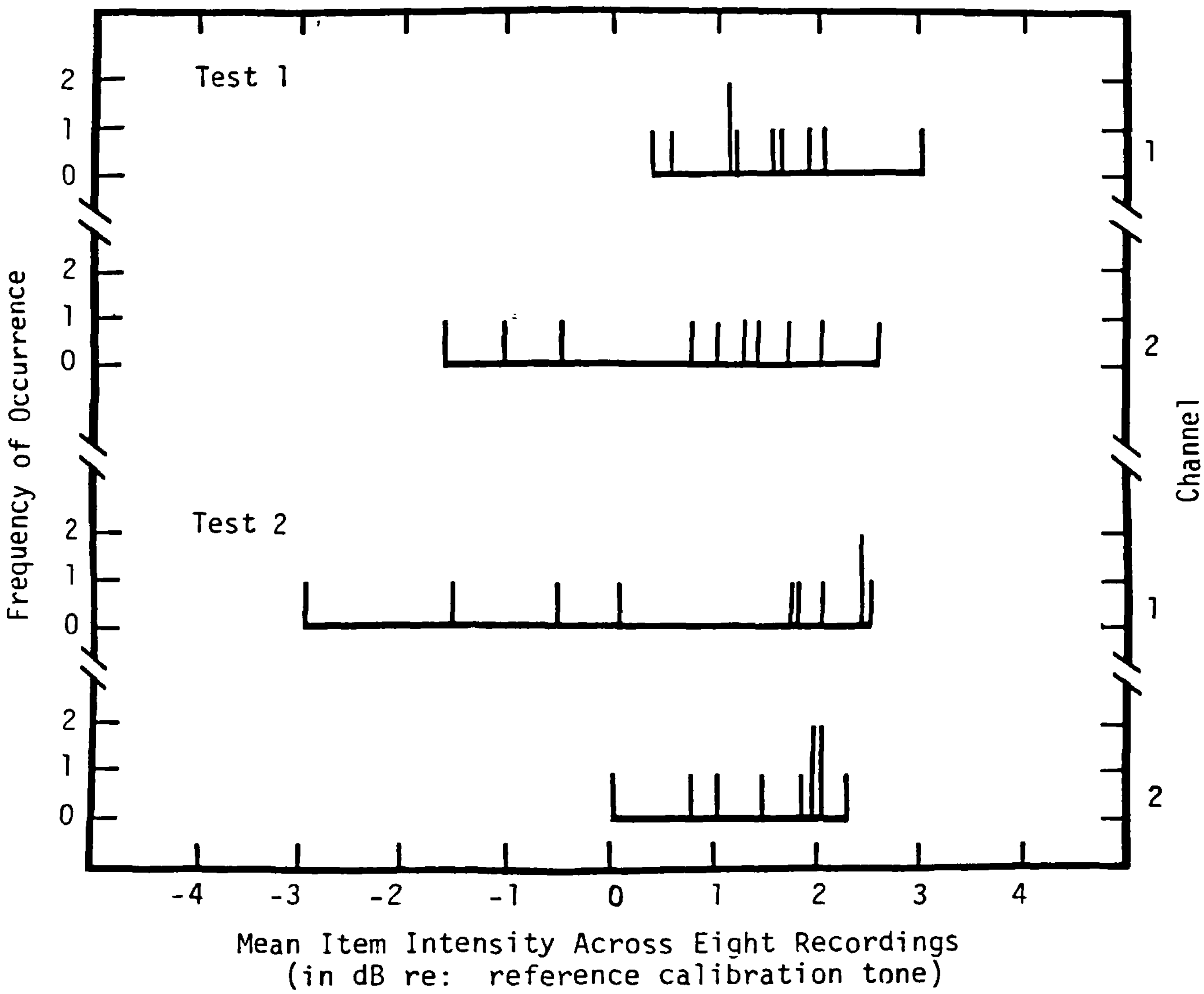

Test 1
Test 2
Frequency of Occurrence
Channel
2
1
0
1
2
-4
-3
-2
-1
0
1
2
3
4
Mean Item Intensity Across Eight Recordings
(in dB re: reference calibration tone)

Figure 1

Table 1 depicts the ranges of measured intensity levels between tapes as well as within tapes. A range of 27.5 dB between the lowest intensity level (-12.0 dB - Tape #1) and the highest (15.5 dB - Tape #3) between tapes (re: the reference calibration tone) was noted. The range of item-intensity levels within tapes varied from a minimum of 7.0 dB (Tape #2) to a maximum of 13.5 dB (Tape #4).

In addition, on three of the eight test tapes used in the present experiment, a severe error in recording technique was demonstrated for item 25 (for both channels). Specifically, the last phoneme of the last word was omitted rendering that word completely unintelligible.

LPFS

Table 2 depicts the ranges of measured intensity levels between and within tapes. Between-tape measures exhibited a range of 32.0 dB - the lowest intensity level being -17.0 dB (Tape #7) and the highest, 15.0 dB (Tape #4). Overall within-tape differences ranged from 7.5 dB (Tape #1) to 23.0 dB (Tape #4).

TABLE 1

CST

Within and between tape comparison
of intensity ranges

TAPE	RANGE*		DIFFERENCES
	Low	High	
#1	-12.0	-0.5	11.5
#2	-3.0	4.0	7.0
#3	2.5	15.5	13.0
#4	-4.5	9.0	13.5
#5	-10.0	2.0	12.0
#6	-8.0	2.0	10.0
#7	-7.5	1.0	8.5
#8	1.0	9.0	8.0

*: Range is in dB re:
reference calibration tone.

TABLE 2

LPFS

Within and between tape comparison
of intensity ranges

TAPE	RANGE*		DIFFERENCES
	Low	High	
#1	-13.5	-6.0	7.5
#2	-6.5	1.5	8.0
#3	-7.5	6.0	13.5
#4	-8.0	15.0	23.0
#5	-10.0	-1.0	9.0
#6	-15.0	-6.5	8.5
#7	-17.0	-8.5	8.5
#8	-7.0	13.0	20.0

*: Range is in dB re:
reference calibration tone.

BF

Table 3 reveals a range in intensity levels between tapes for the 40 syllables in the order of 60.0 dB with the lowest recorded syllable intensity level measured at 39.0 dB below the reference calibration tone (Tape #1) and the highest measured at 21.0 dB above the reference tone (Tape #8). The range of syllable intensity levels within a given tape ranged from an overall difference of 27.5 dB (Tape #7) to 53.5 dB (Tape #4).

To further investigate syllable and word discrepancies the average difference between low-pass and high-pass conditions was computed for each syllable as well as for the spondaic word as a whole. Table 4 reveals that the least discrepant range in intensity levels for the two syllables of a word (across filter conditions) was 2.0 dB (Tape #7) with the most discrepant range between two syllables for a given word being 49.5 dB (Tape #4). Within a single tape, the intensity ranges of syllables within a given word (difference between the most discrepant and the least discrepant syllables within a word) varied from a minimum of 16.0 dB (Tape #3) to a maximum of 35.0 dB (Tape #4).

TABLE 3

BF

Within and between tape comparison
of intensity ranges

TAPE	RANGE*		DIFFERENCES
	Low	High	
#1	-39.0	-5.5	33.5
#2	-35.0	4.0	39.0
#3	-29.0	4.5	33.5
#4	-34.0	19.5	53.5
#5	-34.0	1.0	35.0
#6	-32.5	3.5	36.0
#7	-33.0	-5.5	27.5
#8	-28.5	21.0	49.5

*: Range is in dB re:
reference calibration tone.

TABLE 4

BF

Within and between tape comparison
of syllable and word discrepancies

TAPE	SYLLABLE*		WORD*	
	LD**	MD**	LD**	MD**
#1	6.0	31.5	12.5	30.8
#2	15.0	36.0	19.0	35.2
#3	16.5	32.5	20.3	31.0
#4	14.5	49.5	20.7	48.3
#5	14.0	33.0	18.5	31.7
#6	10.0	30.0	18.8	29.0
#7	2.0	25.5	6.5	22.8
#8	16.0	45.5	18.5	33.5

*: Difference between low-pass and high-pass conditions.

**: LD: least discrepant.
MD: most discrepant.

Word discrepancies were obtained by averaging the two-syllable discrepancy for each word. Thus, the least discrepancy for any given word, between low-pass and high-pass (averaged across the two syllables), was 6.5 dB (Tape #7) whereas the greatest discrepancy was 48.3 dB (Tape #4). Within tapes, word-level discrepancies varied within a range of 15 dB (on the average) from the least discrepant to the most discrepant (the actual range of discrepancies was from 10.2 dB to 18.3 dB).

For all tapes, items recorded on the high band-pass channel tended to be recorded within 5.0 dB of the noise floor of the recording instrument if not actually in the noise floor, as was frequently the case.

RASP

Table 5 shows the discrepancies between items within and between tapes. The range of discrepancies for the test items recorded intensity levels was from -18.0 dB (Tape #1) to 14.0 dB (Tape #3). Within tapes, differences in recorded intensity levels ranged from 7.5 dB (Tape #8) to 17.0 dB (Tape #7).

TABLE 5

RASP

Within and between tape comparison
of intensity ranges

TAPE	RANGE*		DIFFERENCES
	Low	High	
#1	-18.0	-1.5	16.8
#2	-1.5	11.0	12.5
#3	2.0	14.0	12.0
#4	0.5	9.0	8.5
#5	-4.0	6.0	10.0
#6	-7.0	3.5	10.5
#7	-8.0	9.0	17.0
#8	0.0	7.5	7.5

*: Range is in dB re:
reference calibration tone.

DISCUSSION

Willeford's initial concern in developing the CAP battery was to provide a means of evaluating the nature of central-auditory processing problems in children. The use of recorded materials allows the examiner to control such variables as presentation levels and environmental conditions. Therefore, the usage of such carefully controlled recorded materials should allow for the replication by various researchers of normative data, contingent on strict adherence to the procedural protocols as dictated by Willeford. Although Willeford has claimed that the tape recordings of the CAP test battery were experimental, their widespread dissemination throughout the nation (for a fee) would indicate that they are being used as part of routine clinical armamentaria rather than for strictly research purposes.

That various researchers have attempted and failed to replicate the normative data published by Willeford (White, 1977; Mittenberger <u>et al.</u>, 1979), provides a basis for a healthy skepticism concerning the reliability of the published normative data. For example, normative data obtained on children between the ages of 6 to 10 years, have revealed scores as low as 15% on the BF test (White, 1977). Willeford (1978), however, describes his normative data as resulting in a low score of 55%. The rather unacceptably large variability in scores, as reported by these two investigators, probably obviates the clinical usefulness of the test battery in the differentiation of performance scores consistent with pathological conditions from performance scores consistent with normal auditory function (Plakke, 1978).

While Willeford (1977) acknowledges that the ranges of normal-performance scores he has obtained are rather wide - which he views as a factor "probably influenced heavily by the delicate acoustic properties of the stimuli" - he points out that the BF test gives "trouble" on "some" equipment thus providing a rationale for the significantly wider ranges obtained by various investigators in their collection of normative data (White, 1977; Mittenberger, et al., 1978). He states that inaccurate pure-tone measurements, dirty or misaligned tape-recorder playback heads, lack of constant tape speed (of 7-1/2 ips) thus possibly creating a "wow" or "flutter", and use of 1/4 track rather than 1/2 track playback units thus reducing the overall output SPL of the tape, when taken together, account for the variability of data reported in published research. He further contends that "copies" of tapes may incur dubbing errors and would also be of reduced quality. Plakke (1978) further cautioned regarding the use of dubbed tapes as the signal-to-noise (S/N) ratio would be reduced. If difficulties with calibration are insufficient to account for noted discrepancies, Willeford has provided another point of information that may affect the magnitude of the discrepancies found. Specifically, he has asserted that the absence of any culturally deprived areas in Fort Collins, such as those found in many large urban areas may have biased his data.

Yet, findings from the present study may be used to support an alternate explanation for the observed variability in the published normative data obtained by the various researchers. While mean intensity levels of as much as 10.0 dB, as noted on three of the four tests - CST, LPFS, and RASP - may be of questionable clinical relevance, statistical differences, significant at the 0.01 level of significance, were apparent both within and between tapes on each of the measured tests. However, when measured intensity ranges (rather than mean intensity levels) are considered, significant differences between tapes are found, thus precluding the rational use of two different tapes for any given patient. For example, measurements of the CST on

Tape 1 produced values ranging from -12.0 dB to -0.5 dB (a range of 11.5 dB). While the range of values for Tape 3 (13.0 dB) was comparable to that of Tape 1, these values ranged from 2.5 dB to 15.5 dB. Differences between the two minima and the two maxima for these tapes were in the order of 15 dB. In addition, the range of recording levels for the items would suggest the possibility of overdriving some instrumentation when the reference calibration tone is used to calibrate the output of the instruments with a VU-meter reading of zero (since some items then peg the meter). Hence, the use of two different tapes in repeated testing with a patient may result in erroneous interpretation unless differences in recording levels are taken into account. The only possible manner for these corrections to be applied scientifically, is via the generation of performance-intensity (P-I) functions. Since these data are unavailable, conclusions about changes in performance of patients based on different tapes may be invalid, and may result in professional malpractice.

While some clinicians may consider that within-tape mean differences of less than 10 dB are not significant clinically for CST, LPFS, and RASP, the greater discrepancies noted for the BF test cannot be overlooked. It is frustrating to note that while the BF test has been reputed to be one of the more diagnostically significant tests of the CAP battery (White, 1977; Willeford, 1976), it is also the most poorly recorded in the present sample. In test presentation, the sensation level is referenced to thresholds obtained with presumably correctly calibrated materials. If the BF test is recorded such that the material is 39.0 dB less than the reference calibration tone (as in Tape #1), then the materials nominally presented at 30-dB SL (as per Willeford's test instructions) are in fact presented at a level of -9.0 dB SL (9.0 dB below the patient's threshold). Hence, it should not be surprising to find that children with learning disabilities do poorly on this test. One should not be surprised to find that many patients should do poorly.

In an apparent attempt to appease criticisms emerging from the observed excessive variability in scores obtained on the BF test, Willeford simplistically recommends an alternate procedure: one band (either one) is presented at a level of 70- to 80-dB SL. If difficulty persists the two bands are then played, often just slightly below the 80-dB SL ceiling. Clearly performance at 70-dB SL will differ from that at 30-dB SL. The usage of greater sensation levels should be preceded by the acquisition of normative data since none exist or are published at the present time. In terms of findings from this study, the use of the high-pass channel - of significantly reduced intensity - is questionable particularily when most of the stimuli have been recorded within the noise floor of the recording instrument. Presentation of this distorted material at high sensation levels would not be expected to ease the task, since an increased presentation level will not alter the S/N ratio. Findings have indicated an improvement with the increased SL; however, these individuals still remained below the 75% criterion for normal.

Finally, Willeford has recommended that facilities obtain their own normative data. A clinical consideration that is affected by the need for each individual clinic to generate its own data (since one set of norms may not be comparable to another set of norms) is the design of therapy. Indeed, it would be difficult to devise appropriate therapeutic techniques for a child who, having been identified as abnormal in one clinic, becomes normal at a different clinic. One would expect (from a valid test) that a child identified as abnormal in one clinic, also would be identified as abnormal at a different clinic.

In conclusion, the findings of the present study indicate that the quality control associated with Willeford's recordings of his CAP battery is inconsistent. Thus, the usage of his test battery on patient populations MUST be preceded by the acquisition of normative data on a substantial number of subjects matched by sex, age, and other relevant factors to each patient population (a suggestion wisely recommended by Willeford). Since there are 200 test items (across the four tests), a minimum number of subjects (one per item) would be 200 for each matched group. In addition, if item, or regression, or correlation analyses are desired, a minimum of 10 subjects per item is warranted (Nunally, 1978). Under these conditions, at least 2000 subjects would be necessitated. Clearly, this tedious procedure would not be cost efficient, and few clinics would have the luxury to obtain the minimum data necessary. However, it is suggested that, because of its hypothetical diagnostic usefulness, more data using more carefully controlled recorded materials, be initiated and pursued with vigor and independently by various interested researchers.

REFERENCES

AMERICAN NATIONAL STANDARDS INSTITUTE. Specifications for audiometers (S3.6-1969). New York, New York:Acoustical Society of America, 1969.

ARNOLD, H.D. In H. Fletcher, Speech and hearing in communication. Princeton, NJ:D. Van Norstrand Co, Inc, 1953, 216.

BERLIN, C.I. New developments in evaluating central auditory mechanisms. Annals of Otology Rhinology and Laryngology, 1976, 85, 833-841.

BERLIN, C.I. Personal communication at the Annual Meeting of the Canadian Speech & Hearing Association, Edmonton, Alberta, 1981.

BERLIN, C.I. & LOWE, S.S. Temporal and dichotic factors in central auditory testing. In J. Katz (Ed.) Handbook of clinical audiology. Baltimore, MD:Williams & Wilkins, 1972.

BERLIN, C.I. & McNEILL, M.R. Dichotic listening. In N. J. Lass (Ed.) Contemporary issues in experimental phonetics. New York:Academic Press, 1976, 327-387.

BERLIN, C.I., HUGHES, L.F., LOWE-BELL, S.S. & BERLIN, H.L. Dichotic right ear advantage in children 5 to 13. Cortex, 1973, 9, 394-402.

BERLIN, C.I., LOWE-BELL, S.S., JANETTA, P.J. & KLINE, D.G. Central auditory deficits after temporal lobectomy. Archives of Otolaryngology, 1972, 96, 5-10.

BILLINGSLEA, B. Intelligibility of distorted speech. Master's thesis, Colorado State University, Fort Collins, CO, 1963.

BOCCA, E. Clinical aspects of cortical deafness. Laryngoscope, 1958, 68, 301-309.

BOCCA, E. & CALEARO, C. Central hearing processes. In J. Jerger (Ed.) Modern developments in audiology (1st. edition). New York:Academic Press, 1963, 337-370.

BOCCA, E., CALEARO, C. & CASSINARI, V. A new method for testing hearing in temporal lobe tumors; Preliminary report. Acta Otolaryngologica, 1954, 44, 219-221.

BOCCA, E., CALEARO, C., CASSINARI, V. & MIGLIAVACCA, F. Testing "cortical" hearing in temporal lobe tumors. Acta Otolaryngologica, 1955, 45, 289-304.

BROADBENT, D.E. The role of auditory localization in attention and memory span. Journal of Experimental Psychology, 1954, 47, 191-196.

BROADBENT, D.E. Attention and the perception of speech. Scientific American, 1962, 206, 143-151.

BROADBENT, D.E. & GREGORY, M. Accuracy of recognition for speech presented to the right and left ears. Quarterly Journal of Experimental Psychology, 1964, 16, 359-360.

BRUNT, M. & GOETZINGER, C.P. A study of three tests of central function with normal hearing subjects. Cortex, 1968, 4, 288-297.

CALEARO, C., & ANTONELLI, R. "Cortical" hearing tests and cerebral dominance. Acta Otolaryngologica, 1963, 56, 17-26.

CALEARO, C. & ANTONELLI, A.R. Audiometric findings in brainstem lesions. Acta Otolaryngologica, 1968, 66, 305-319.

DEMPSEY, C. Some thoughts concerning alternate explanations of central auditory test results. In R.W. Keith (Ed.) Central auditory dysfunction. New York:Grune & Stratton, Inc., 1977, 293-317.

DERMODY, P. & HOFFSINGER, P.D. Auditory processing factors in dichotic listening. Paper presented at the Annual Meeting of the Acoustical Society of America, Washington, D.C., 1976.

GAMBINO, J. Filtered-speech articulation curves. Master's thesis, Colorado State University, Fort Collins, CO, 1963.

HAYASHI, R., OHTA, F., & MORIMOTO, M. Binaural fusion test; A diagnostic approach to the central auditory disorders. International Audiology, 1966, 5, 133-135.

IVEY, R.G. Tests of CNS auditory function. Master's thesis, Colorado State University, Fort Collins, CO, 1969.

JERGER, J. Audiological manifestations of lesions in the auditory nervous system. Laryngoscope, 1960, 70, 417-425.

JERGER, J. Diagnostic audiometry. In J. Jerger (Ed.) Modern Developments in Audiology (2d. edition). New York:Academic Press, 1973, 75-116.

JERGER, J. & JERGER, S. Auditory findings in brainstem disorders. Archives of Otolaryngology, 1974, 99, 342-350.

JERGER, J., SPEAKS, C. & TRAMMELL, J.A. A new approach to speech audiometry. Journal of Speech and Hearing Disorders, 1968, 33, 318-328.

JERGER, S. & JERGER, J. Extra- and intra-axial brainstem auditory disorders. Audiology, 1975, 14, 93-117.

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

KATZ, J., BASIL, R.A. & SMITH, J.M. A staggered spondaic word test for detecting central auditory lesions. Annals of Otology Rhinology and Laryngology, 1963, 72, 908-917.

KEITH, R. Synthetic sentence identification. In R.W. Keith (Ed.) Central auditory dysfunction. New York:Grune & Stratton, Inc., 1977, 73-102.

KIMURA, D. Cerebral dominance and the perception of verbal stimuli. Canadian Journal of Psychology, 1961, 15, 166-171.

LARIMORE, H.W. & WILLEFORD, J.A. Central auditory processing in stutterers. Paper presented at the Annual Meeting of the American Speech and Hearing Association, Detroit, MI, 1980.

LINDEN, A. Distorted speech and binaural speech resynthesis tests. Acta Otolaryngologica, 1964, 58, 32-48.

LOGAN, H.A. Performance of three and four year-old children on the Alternating Speech Perception Test. Master's thesis, Emerson State College, 1979.

LYNN, G. & GILROY, J. Neuro-audiological abnormalities in patients with temporal lobe tumors. Journal of Neurological Sciences, 1972, 17, 167-184.

LYNN, G. & GILROY, J. Evaluation of central auditory dysfunction in patients with neurological disorders. In R.W. Keith (Ed.) Central auditory dysfunction. New York:Grune & Stratton, Inc., 1977, 177-222.

LYNN, G, BENITEZ, J., EISENBREY, A., GILROY, J. & WILNER, H. Neuro-audiological correlates in cerebral hemisphere lesions: Temporal and parietal lobe tumors. Audiology, 1972, 11, 115-134.

MATZKER, J. Two new methods for the assessment of central auditory functions in cases of brain disease. Annals of Otology Rhinology and Laryngology, 1959, 68, 1185-1197.

MITTENBERGER, G., DAWSON, G. & RAICA, T. Central auditory testing with peripheral hearing loss. Archives of Otolaryngology, 1978, 104, 11-15.

MITTENBERGER, G., CARUSO, V., CORREIA, M., LOVE, T. & WINKLEMAN, P. Utilization of a central auditory processing test battery in diagnosing decompression sickness. Journal of Speech and Hearing Disorders, 1979, 44, 111-120.

MUSIEK, F. & GERURKIRK, N. Auditory perceptual problems in children: Considerations for the otolaryngologist and audiologist. Laryngoscope, 1980, 90, 962-971.

N.A.B. Standard magnetic tape recording and reproducing (reel-to-reel). Washington, DC:National Association of Broadcasters, Engineering Dept., 1965.

NOFFSINGER, P.D. & KURDZIEL, S.A. Assessment of central auditory lesions. In W.F. Rintelmann (Ed.) Hearing assessment. Baltimore, MD:University Park Press, 1979, 351-377.

NOFFSINGER, P.D., OLSEN, W.O., CARHART, R., HART, C.W. & SAHGAL, V. Auditory and vestibular aberrations in multiple sclerosis. Acta Otolaryngologica (suppl.), 1972, 303, 1-63.

NUNALLY, J. Psychometric theory. New York:McGraw-Hill, 1978.

OHTA, F., HAYASHI, R. & MORIMOTO, M. Differential diagnosis of retrocochlear deafness: Binaural fusion tests and binaural separation test. International Audiology, 1967, 6, 58-62.

PALVA, A. Filtered speech audiometry. Basic studies with Finnish speech towards the creation of a method for the diagnosis of central hearing disorders. Acta Otolaryngologica (suppl.), 1965, 210, 1-86.

PALVA, A. & JOKINEN, K. The role of the binaural test in filtered speech audiometry. Acta Otolaryngologica, 1975, 79, 310-314.

PETERSON, G. & LEHISTE, I. Revised CNC lists for auditory tests. Journal of Speech and Hearing Disorders, 1962, 27, 62-70.

PINHEIRO, M. Tests of central auditory function in children with learning disabilities. In R.W. Keith (Ed.) Central auditory dysfunction. New York:Grune & Stratton, Inc., 1977, 223-256.

PLAKKE, B.L. Children's performance on a binaural fusion task. Doctoral dissertation, Memphis State University, 1978.

SCHULTZ, C.H. Central auditory function in stuttering and non-stuttering children. Master's thesis, Central Michigan University, 1977.

SHANKWEILER, D. Effects of temporal-lobe damage on perception of dichotically presented melodies. Journal of Comparative Physiological Psychology, 1966, 62, 115-119.

SHANKWEILER, D. & STUDDERT-KENNEDY, M. Identification of consonants and vowels presented to left and right ears. Quarterly Journal of Experimental Psychology, 1967, 19, 59-63.

SMITH, B.B. & RESNICK, D.M. An auditory test for assessing brainstem integrity: Preliminary report. Laryngoscope, 1972, 82, 414-424.

SPEAKS, C. Dichotic listening; A clinical or research tool? In M. Sullivan (Ed.) Central auditory processing disorders. Omaha:University of Nebraska Press, 1975.

SPEAKS, C. & JERGER, J. Method for measurement of speech identification. Journal of Speech & Hearing Research, 1965, 8, 185-194.

TOBEY, E.A., CULLEN, J.K., Jr. & RAMPP, D.L. Effects of stimulus-onset asynchrony on the dichotic performance of children with auditory-processing disorders. Journal of Speech & Hearing Research, 1979, 22, 197-211.

TREISMAN, A.M. The effect of irrelevant material on the efficiency of selective listening. American Journal of Psychology, 1964, 77, 533-546.

WELSH, L., WELSH, J. & HEALY, M. Central auditory testing and dyslexia. Laryngoscope, 1980, 90, 972-984.

WHITE, E. Children's performance on the SSW test and Willeford battery: Interim clinical data. In R.W. Keith (Ed.) Central auditory dysfunction. New York:Grune & Stratton, Inc., 1977, 319-340.

WILLEFORD, J. Central auditory function in children with learning disabilities. Paper presented at the Annual Meeting of the American Speech & Hearing Association, Las Vegas, NV, 1974.

WILLEFORD, J. Central auditory function in children with learning disabilities. Audiology & Hearing Education, 1976, 2, 12-20.

WILLEFORD, J. Assessing central auditory behavior in children: A test battery approach. In R.W. Keith (Ed.) Central auditory dysfunction. New York:Grune & Stratton, Inc., 1977, 43-72.

WILLEFORD, J. Sentence tests of central auditory function. In J. Katz (Ed.) Handbook of clinical audiology (2d. edition). Baltimore:Williams & Wilkins, 1978, 252-261.

WINER, B.J. Statistical principles in experimental design (2d. edition). New York:McGraw-Hill, 1971.

WINKELAAR, R.G. & LEWIS, T.K. Audiological tests for evaluation of central auditory disorders. Journal of Otolaryngology, 1976, 6, 127-134.

APPENDIX A

UNDERLYING BASIS OF EACH TEST OF THE C.A.P. TEST BATTERY

The CAP test battery is composed of four tests designed to assess the integrity of the auditory system at various levels of the CANS, from the brainstem -- Binaural Fusion (BF) and Rapidly Alternating Speech Perception (RASP); to thalamo-cortical levels and the auditory cortex -- Low-Pass Filtered Speech (LPFS); through to the temporal lobes -- Competing Sentence Test (CST).

The discussion that follows will analyze each of the four subtests of the CAP test battery in terms of previous research, including developments that led to the manufacture of each test, Willeford´s recommended procedures, a listing of the test items, and a summary of the current clinical findings associated with each test.

1. BF

Review of Previous Research

Arnold (cited by Fletcher, 1953) was one of the first individuals to demonstrate the capability of the brain to receive and integrate two differentially filtered speech sounds. These sounds are presented simultaneously, with a low-pass filtered signal to one ear, and the same signal high-pass filtered to the other ear (with a cut-off frequency of 1000 Hz). Arnold was able to demonstrate conclusively enhanced discrimination scores for this binaural condition (when compared to discrimination scores for either signal presented monaurally).

Matzker (1959) adapted Arnold's procedure for use in the assessment of brainstem integrity. Matzker employed 41 two-syllable German PB words, which were band-pass filtered between 500 and 800 Hz in one channel and between 1815 and 2500 Hz in the other channel. The words thus filtered on each channel were unintelligible. However, when the low band-pass filtered signal is presented to one ear and the high band-pass signal is presented to the other ear simultaneously, intelligibility is enhanced. Words, presented at a sensation level (SL) previously determined to result in maximum discrimination performance scores, were presented in three varying conditions: (1) binaural - the simultaneous presentation of the low band-pass filtered stimulus to the left ear and the high band-pass filtered stimulus to the right ear; (2) diotic - the simultaneous presentation of both low and high band-pass filtered stimuli to both ears; and (3) binaural - repetition of condition (1). Matzker viewed the integration of the two filtered stimuli as taking place within the brainstem, probably at the level of the cochlear nuclei and medial geniculate bodies. He postulated that failure of the integration was indicative of a loss of ganglionic cells within the primary auditory centers. Thus, normal brainstem function was indicated by the demonstration of greater performance scores in the diotic condition than in either binaural condition. In addition, he noted that fewer mistakes were obtained on the second binaural trial than on the first. Brainstem involvement was indicated by increased number of errors, particularly in the two binaural conditions relative to errors committed under the diotic condition. Matzker also reported post-mortem histopathologic findings that supported his claim that binaural integration occurs at brainstem levels.

The effectiveness of Matzker's test has been challenged by Linden (1964) after he failed to replicate Matzker's results. However, Linden, a priori, changed the filter characteristics, test protocols, and interpretation procedures from those recommended by Matzker. In spite of these significant changes, Linden attributed differences between his results and those of Matzker as indications of errors in Matzker's technique and interpretations. In Linden's test, 15 lists of nine spondee words each were band-pass filtered between 560 and 715 Hz, and 1800 and 2200 Hz, and were presented under varying conditions. These conditions included: (1) monaural - through the low-pass filter; (2) monaural - through the high-pass filter; (3) monaural - through both low- and high-pass filters simultaneously; (4) binaural - through low-pass filter to one ear and high-pass filter to the other ear. He recommended a presentation level of less than 40-dB SL because he noted differential effects of presentation level as a function of filter characteristics. Specifically, he observed that performance scores improved considerably for low band-pass stimuli as the sensation level was increased, while performance scores improved only marginally with increased sensation level for high band-pass stimuli. He reported results obtained on five cases of confirmed temporal-lobe damage. Of these, there was a reduction in the monaural condition for the performance score of the ear contralateral to the cortical lesion in every case. In addition, reduced scores were obtained in the binaural condition for three of the cases, and one case yielded equally reduced performance scores under all conditions. Linden also reported data obtained on 13 patients with expanding intracranial lesions outside of the temporal lobe. Ten of these yielded normal intelligibility scores under the binaural condition, while two cases showed equally reduced intelligibility scores across all conditions, and one case resulted in a reduced monaural performance score for both ears while, at the same time demonstrating normal performance scores for the binaural condition. Despite the occurence of reduced intelligibility in the binaural condition for six of the 18 reported cases, Linden scored all the results as being negative since he did not view the binaural condition as being significantly worse, thus flagrantly disregarding the fact that three of these six cases showed better monaural scores than binaural scores (Lynn and Gilroy, 1972). Linden concluded that Matzker's test was not a useful measure for the assessment of CANS function. Lynn and Gilroy (1972) cautioned that Linden failed to provide documentation of the presence of brainstem lesions in his cases, and that perhaps they had only cerebral involvement.

Further criticism of Matzker's procedure was initiated by Hayashi, <u>et al.</u> (1966). These investigators noted reduced binaural scores in 10% (3/30) of their cases of undefined retrocochlear pathology. Modifications of Matzker's technique, to include six presentation conditions of monosyllabic nonsense syllables filtered through frequency bands of 300 to 600 Hz, and 1200 to 2400 Hz (they claimed that Matzker's high band-pass was 1500 to 2400 Hz) also implemented the following changes: (1) dichotic presentation of non-distorted syllables (a diotic condition); (2) dichotic presentation of the signals having passed through the high and low band-pass filters to both ears (another diotic condition); (3) dichotic presentation - high band-pass to the left ear, low band-pass to the right ear; (4) monotic presentation of the high band-pass signal to the left ear; (5) dichotic presentation - low band-pass to the left ear, and high band-pass to the right ear; (6) monotic presentation of the high band-pass signal to the right ear. These test conditions were administered to 78 patients with varying cortical pathologies. Reduced binaural-function scores were reported in five cases (including temporal-lobe epilepsy, cerebellar ataxia, sensorineural hearing loss, and two cases with head trauma). Normal pattern of binaural fusion (undefined, but presumably based upon the performance on dichotic conditions) was apparent in all cases of Meniere's disease and of acoustic trauma. Thus, they concluded that performance scores for the binaural-fusion test remain unaffected by cochlear lesions. They further reported a "tendency" of poor binaural function when the high band-pass signal is presented to the ear contralateral to the affected cerebral area. On the basis of these data, they concluded that binaural fusion would be affected detrimentally by cortical lesions, as well as brainstem lesions.

In a further study (Ohta, et al., 1967), the same laboratory reported results from 130 cases of "perceptual" (sic) deafness. They found that of 27 cases of temporal head injury, seven cases demonstrated poor binaural fusion. In addition, both cases of parietal, and neither case of frontal injury, showed abnormal auditory fusion, while only one of four occipital-injury cases resulted in abnormal binaural fusion. They re-emphasized that binaural fusion was a function of cortical function, and not restricted to brainstem function. Criticisms of their conclusions were advanced by Plakke (1978) who noted that the failure to define the subject populations in both studies provided inconclusive evidence of the presence or absence of brainstem pathology from which they drew their conclusions. The criterion used by Hayashi, <u>et al.</u> (1966), for normal performance was based on a limited sample size (N = 10). The criterion to which Hayashi and colleagues adherred was not explicitly stated. In addition, as noted by Plakke (1978), it is possible that the high band-pass filter was wide enough that binaural fusion became unecessary to obtain adequate performance scores.

Matzker's procedure did receive support from Smith and Resnick (1972). These investigators reported findings in agreement with those published by Matzker, on their test of dichotic binaural fusion. This test was designed to differentiate brainstem lesions from end-organ or temporal-lobe pathologies. Test stimuli consisted of three lists of 50 revised CNC words (Peterson and Lehiste, 1962), which were band-pass filtered (360 to 890 Hz, and 1750 to 2220 Hz) with the peak amplitude of the high band-pass signal being 10 dB greater than the amplitude of the low band-pass signal. The test consisted of three binaural conditions: (1) dichotic "A" - high band-pass signal to the left ear, low band-pass to the right ear; (2) diotic - both band-pass bands to both ears; (3) dichotic "B" - reversal of ears from the dichotic "A" condition. Positive test findings, indicated by dichotic enhancement of performance scores, were reported in cases of confirmed brainstem lesions. Negative findings, suggestive of normal brainstem function, were reported in patients with normal hearing, bilateral peripheral hearing loss and documented cases of temporal-lobe lesions.

Hoffsinger, <u>et al.</u> (1972), adapted Smith and Resnick's technique for use with NU #6 materials. Words were band-pass filtered between 250 and 700 Hz and between 1250 and 1750 Hz. They added two monaural conditions (each ear receiving both high band-pass and low band-pass signals monaurally) to those of Smith and Resnick. Test interpretation was undertaken separately for each ear, and scores of less than 60% were viewed as abnormal independent of stimulus presentation paradigm. Administration of this test to a group of multiple-sclerotic patients revealed abnormal findings in five of 36 reported cases.

Palva and Jokinen (1975) in a modification of the procedure advocated by Smith and Resnick, changed the band-pass filter settings to 480 to 720 Hz and 1820 to 2400 Hz. In addition, they automated the presentation paradigm and adhered to the following order: (1) word #1 - presentation of both band-passed signals to the right ear; (2) word #2 - presentation of both band-passed signals to the left ear; (3) word #3 - presentation of the high band-pass to one ear and the low band-pss to the other ear; (4) word #4 - reversal of condition 3. This order was repeated until all 90 test items had been presented. Performance scores were obtained for the right ear, the left ear, and the binaural stimulation conditions. The most common positive test finding was an asymmetry in discrimination scores between ears as noted in cases of presbycusis, peripheral hearing losses, and central lesions (intracranial tumors, multiple sclerosis, and skull trauma). According to Palva and Jokinen (1975), positive binaural findings, while noted in several of these cases, were most evident in those patients suffering from intracranial vascular disorders or head trauma. Palva (1965) cautioned against interpreting the positive binaural score as indicative of a brainstem lesion in the presence of a positive monaural score. He asserted that a peripheral lesion would result in the deterioration of the binaural score, while a higher level lesion would not affect this score, in the presence of decreased monaural scores.

Ivey (1969) developed the BF which Willeford now uses in his CAP test battery. For this test, spondee words were selected on the basis of their relative unintelligibility when presented through either band-pass filter, while maintaining high intelligibility when presented through both band-pass filters simultaneously.

> "In order to achieve the desired filtering characteristics the following procedure was used. The original tape was placed on the PR-10 tape recorder. Channel A" [containing the carrier phrase] "was passed through a mixer (unfiltered) and recorded on channels A and B of the Ampex 354. Channel B" [of the PR-10] "was passed through an attenuator, in order to make the presentation level of the words more uniform, then the unfiltered signal was routed in two ways: 1) through a 700 Hz low-pass filter, with a 36 dB/octave slope, and recorded on Channel A, and; 2) through a 1900 Hz high-pass filter, with a slope of 36 dB/octave and recorded on Channel B of the 354 recorder. These rejection rates were achieved by cascading the 18 dB slope components of two Spencer-Kennedy filters in the appropriate low- or high-pass condition (see Figure 3)" [Figure B1].
>
> "To complete the filtering of the spondee words on each channel without filtering the carrier phrase, the tape resulting from the foregoing process was then placed on the PR-10. Channel A stimuli were passed through a 500 Hz high-pass filter with a 36 dB/octave slope, and then through a mixer to be recorded on Channel A of the 354 recorder. Channel B stimuli were passed through a 2100 Hz low-pass filter with a slope of 36 dB/octave and also passed through the mixer before being recorded on Channel B of the 354. In order to prevent the carrier phrase from being filtered,

> the mixer was used to combine the signals from Channels A and B when the carrier phrase was or (sic) by the manipulation of two silent switches (sic). The result was series of unfiltered carrier phrases on both channels, followed by 500 - 700 Hz band-pass segments of the spondees on Channel A, and 1900 - 2100 Hz band-pass segments of the spondees on Channel B." (Ivey, 1969, pp. 8-10)

A most cursory examination of the block diagram of Ivey's instrumentation for the filtering of the test items reveals some elementary flaws that could have some serious consequences. Specifically, according to the figure he has provided in his thesis, the "A" output of the PR-10 recorder appears to be simply wire-divided to the outputs of the LP and of the HP filters. The absence of a dividing network, since the typical line outputs of tape recorders require load impedances of 10,000 Ohms or more, and the output impedance of Spencer-Kennedy filters is in the order of 600 Ohms, thus creates an impedance mismatch. Thus the tape "A" output appears to be loaded with two 600-Ohm loads in parallel which would yield a load impedance of 300 Ohms. This impedance mismatch, caused by the omission of an impedance-matching dividing network, would cause the current flow to increase by a factor of 33.3.

Figure A1

Block diagram of the instrumentation used by Ivey.

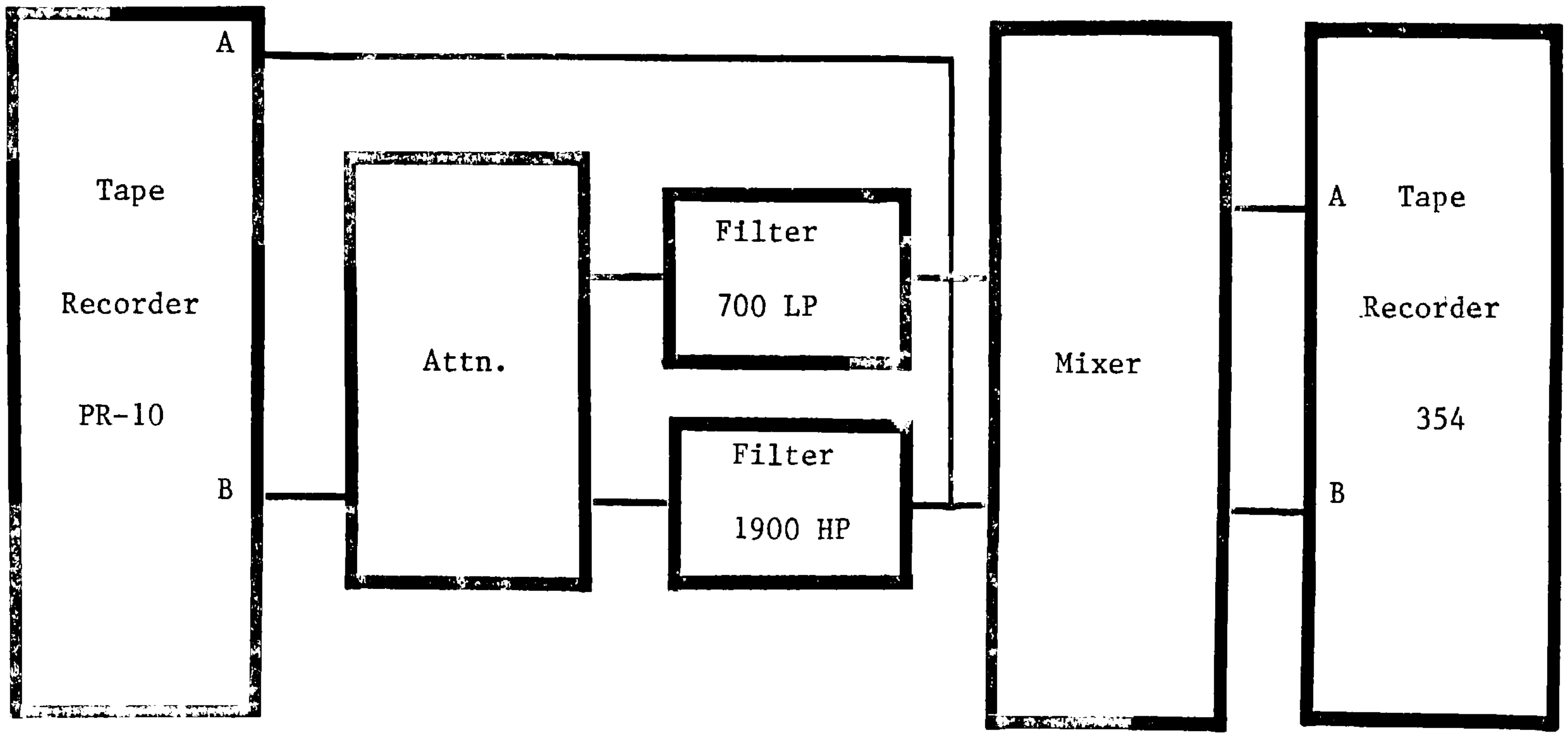

Tape
Recorder
PR-10
A
B
Attn.
Filter
700 LP
Filter
1900 HP
Mixer
A
B
Tape
Recorder
354

Figure A1

Figure A2

Corrected block diagram of the instrumentation used by Ivey.

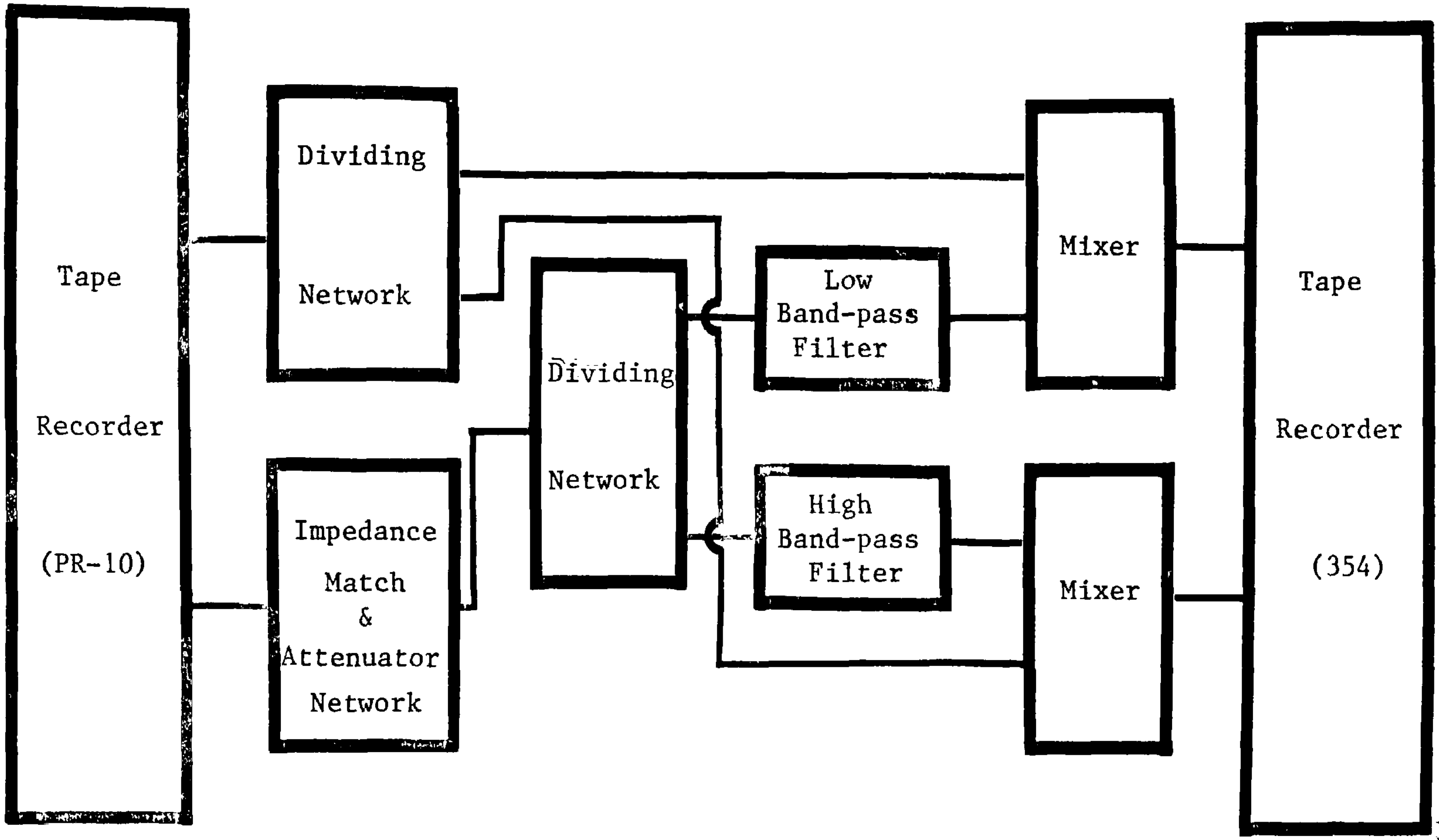

Tape
Recorder
(PR-10)
Dividing
Network
Impedance
Match
&
Attenuator
Network
Dividing
Network
Low
Band-pass
Filter
High
Band-pass
Filter
Mixer
Mixer
Tape
Recorder
(354)

Figure A2

This increased flow of current through the circuit could overload some electronic components and result in a filtering effect of various frequencies, as well as introducing unwanted distortion products in the signal being recorded. In addition, it is unlikely that an attenuator would be provided with one input and two output connections, since the attenuator would then become unbalanced in terms of output impedances. Thus, it is likely that a dividing network was introducec at that point. Furthermore, the input impedances of most attenuators are in the order of 600 Ohms, and without an impedance-matching device, the output impedance of the tape recorder would mismatch that of the input of the attenuator resulting in additional frequency-selective filtering and added distortion products. A block diagram of appropriate instrumentation to accomplish the stated goals of Ivey has been included in Figure A2.

In addition, the text quoted above indicates that the carrier phrase from the "A" output of the PR-10 recorder was fed to a mixer and then to both the "A" and the "B" inputs of the 354 tape recorder. In point of fact, this is a physical impossibility, since a mixer has two or more inputs and only one output. Thus, it is most likely that Ivey fed the "A" output of the PR-10 through a dividing network (one input and two or more outputs), although the inaccuracies of the text and of the figure he provides tend to make some researchers skeptical of Ivey's ability to connect various instruments appropriately. It should be recognized that Ivey's mistakes were approved as consistent with state-of-the-art research by the faculty that constituted his thesis committee of which Willeford was a member. If the "A" output of the PR-10 were wired as shown in Figure B1, then the carrier phrase would undergo some selective filtering effects and distortion caused by the impedance mismatch. If the "B" output of the PR-10 were wired as shown on Figure B1, then the test items would undergo selective filtering both prior to and after the passage through Ivey's sets of filters, and were recorded with attendant distortion products. Ivey's block diagram also would lead one to believe that one mixer was sufficient for the mixing of all signals, which clearly is impossible. In other figures scattered throughout Ivey's thesis, there is never mention of dividing networks, but only of mixers, when most of the needs according to the text required the introduction of dividers. The instrumentation block-diagram shown in Figure B2 was designed to avoid the problems that Ivey may have experienced without realizing. In addition, no documentation of proper tape-recorder calibration, and maintenance was provided by Ivey.

The demonstrated lack of care and attention to proper techniques graphically exemplified by Ivey may have resulted in the production of signals that would not be easily replicable in any other laboratory.

In summary, although general patterns of performance on this task, by patients suffering from various pathologies, have been recognized in the literature, the absence of consistent instrumentation, psychophysical methodology, interpretation criteria and documentation of pathologies may make comparisons across the various researchers tenuous.

Willeford's Recommended Procedures

Two 20-word lists of selected spondee words (Ivey, 1969) filtered into two narrow frequency bands - a low band-pass segment (500 - 700 Hz) and a high band-pass segment (1900 - 2100 Hz) are presented binaurally with one ear receiving the low band, and the other ear the high band. The procedure is repeated with each ear receiving the band-pass segment not previously presented to that ear. Words are presented at 30-dB sensation level (SL) (re: pure-tone threshold for 500 Hz and 2000 Hz for low and high band-pass respectively) (Willeford, 1976). Scoring for a given ear is assigned arbitrarily to the ear receiving the low band-pass segment.

Table A1 contains the test items for the Binaural Fusion test.

TABLE A1

BINAURAL FUSION

List 1

1.	bagpipe	11.	dovetail
2.	woodchuck	12.	shoelace
3.	baseball	13.	bedroom
4.	bloodhound	14.	eyebrow
5.	churchbell	15.	meatball
6.	daylight	16.	bluejay
7.	rainbow	17.	birdnest
8.	drugstore	18.	northwest
9.	bonbon	19.	although
10.	buckwheat	20.	padlock

List 2

1.	doormat	11.	wigwam
2.	footstool	12.	dollhouse
3.	horseshoe	13.	wildcat
4.	stairway	14.	scarecrow
5.	housework	15.	soybean
6.	lifeboat	16.	therefore
7.	mishap	17.	whizbang
8.	nutmeg	18.	workshop
9.	platform	19.	yardstick
10.	watchword	20.	bobwhite

Findings Associated with Willeford's Test

Abnormally reduced scores (less than 75%) have been noted in teenagers and adults suffering from brainstem pathology - which may or may not be secondary to temporal-lobe involvement (Lynn and Gilroy, 1972). In addition, abnormally reduced scores also have been found in patients afflicted with peripheral hearing loss associated with decompression sickness (Mittenberger, et al., 1978). Willeford (1978) and Dempsey (1977) reported that children with auditory problems failed this test more frequently than any of the other subtests of the CAP test battery. Dempsey added that all children identified as severely learning disabled failed either the BF or the LPFS tests. Welsh, et al. (1980), have suggested that the BF test and the LPFS test were the most sensitive tests of the battery for the detection of central-auditory pathology among dyslexic children.

2. RASP

Review of Previous Research

The use of alternating speech as a measure of central auditory function was suggested by Bocca (1960, cited in Bocca and Calearo, 1963), although Calearo (1960, cited in Bocca and

Calearo, 1963) originally had developed the test for the detection of malingerers. The test incorporated sentences which were switched periodically from one ear to the other so that each ear received half of the sentence. The switching rate was varied between 2 and 40 alternations per second. Findings for normal-hearing listeners typically revealed 100% discrimination independent of the switching rate. Lesions confined to the temporal auditory cortex did not appear to affect the accuracy of discrimination. However, in cases of diffuse cerebral pathology and brainstem involvement, discrimination scores were reduced (Bocca and Calearo, 1963). Bocca (1960, cited in Bocca and Calearo, 1963) thus considered the test to be effective in the differential diagnosis of brainstem pathology from peripheral or cortical lesions.

Lynn and Gilroy (cited in Winkelaar and Lewis, 1976) modified Bocca and Calearo's procedure so that the dwell time of the signal was 300 ms. Although the segments, when presented monaurally, were unintelligible, when fused together during binaural presentations, they became readily identified by normal-hearing listeners. Scores of less than 90% (based on 10 test items) were considered to reflect abnormally large amounts of breakdown, and were thought to be related to brainstem pathology. Willeford's recommended procedure virtually is identical to that practiced by Lynn and Gilroy. However, the test sentences used by Willeford are different from those used by Lynn and Gilroy. The effects of that difference remain undocumented.

Willeford's Recommended Procedures

Three lists of 10 sentences each are presented at 30-dB SL with half of the sentences beginning in each ear. The stimuli in each sentence alternate between the two ears every 300 ms.

Table A2 contains the test items for the Rapidly Alternating Speech Perception test.

TABLE A2

RAPIDLY ALTERNATING SPEECH PERCEPTION
(Roman numerals indicate the channel receiving the very first sound)

List 1

1.	II	The fire engine raced down the street.
2.	II	Do you want to go to the picnic?
3.	II	There were many trees around the house.
4.	I	My dog always does what I ask.
5.	I	She is afraid to go home alone.
6.	II	The bird flew out of its cage.
7.	I	The puppy chased the big red ball.
8.	I	The tree branch broke off in the storm.
9.	II	Did you get the tickets for the game?
10.	II	Where did you put the yellow sweater?

List 2

1.	I	We camped in the woods last night.
2.	II	Would you join me for cokes after school?
3.	II	The mayor was elected yesterday.
4.	II	The secretary gave me the wrong number.

TABLE A2 (continued)

5. I There are many kinds of fish in the ocean.

6. II The children enjoyed playing at the beach.

7. I The horse raced around the track.

8. I Put a dozen apples in the sack.

9. I There was dew on the grass this morning.

10. I Plants will begin to grow in the spring.

List 3

1. II He spilled the gravy on the table.

2. II The moon shines brightly in the sky.

3. I The officer gave him a ticket for speeding.

4. I My father takes me fishing every fall.

5. I He fell in the lake and yelled for help.

6. II Did the camera flash scare you?

7. II She carried the parrot on her shoulder.

8. I The garbage man comes on Wednesday.

9. I The bird built a nest in the tree.

10. II I like to drink cocoa for breakfast.

Findings Associated with Willeford's Test

While both the RASP and the BF tests are sensitive to brainstem-level function, collective data (Lynn and Gilroy, 1977) have been interpreted to indicate that the RASP has the added potential of differentiating between higher and lower brainstem disorders. Specifically, researchers, including Lynn and Gilroy (1977), have associated abnormal findings on the RASP with lesions of the caudal regions of the pontine area of the brainstem. Despite the lack of differential sensitivity of the RASP for children afflicted with auditory perceptual disorders (Dempsey, 1977), this test has resulted in abnormal findings for individuals with peripheral hearing loss (Lynn and Gilroy, 1972; Mittenberger, et al., 1978), in individuals following decompression sickness (Mittenberger, et al., 1979), and in a small percentage (approximately 13%) of dyslexic children (Welsh, et al., 1980).

Lynn and Gilroy (1977) cautioned that abnormally reduced performance scores on the RASP test also have been obtained from individuals afflicted only with peripheral hearing impairments. They postulated that abnormal findings in these cases could be related to impaired phonemic discrimination or phonemic regression (resulting from a peripheral sensorineural hearing impairment). In addition, they also noted slightly reduced performance scores in individuals affected by diffuse (bilateral) disorders at the cerebral level (*id est*, neuronal atrophy, with ventricular enlargement resulting in corpus callosum distortion as associated with cases diagnosed as dementia).

3. LPFS

Review of Previous Research

Early work undertaken by Bocca, *et al.* (1954, 1955) led to the development of a procedure for assessing the integrity of the auditory system at the cortical level. Through monaural presentation of low-pass filtered speech, poorer speech-discrimination scores were yielded in the ear contralateral to the cortical lesion. The procedure itself incorporated lists of ten phonetically balanced, disyllabic, meaningful words which were low-pass filtered such that all frequencies greater than 800 Hz practically were eliminated. A series of lists were presented monaurally and an articulation curve plotted for each ear. Discrimination scores of 70 - 80%, bilaterally, were viewed as normal. In cases of temporal-lobe tumors, in the absence of peripheral involvement, discrimination scores were 25% poorer in the ear contralateral to the temporal-lobe tumor (Bocca, 1958). In their 1955 study, Bocca et al. used a low-pass filter with a cut-off frequency of 1000 Hz. The range of articulation scores for their alleged normal listeners increased to 60% to 80%. However, the test appeared sensitive only to lesions confined to the auditory area of the temporal lobe. In addition, lesions outside the auditory area failed to result in significantly decreased performance scores.

Later research (Calearo and Antonelli, 1963) incorporating 32 unilateral temporal-lobe lesion patients, ten with temporal-lobe epilepsy - which involved partial temporal lobectomy including the auditory cortical area - was consistent with earlier findings, with the ear contralateral to the temporal-lobe lesion demonstrating scores that were 20% to 30% reduced compared to the ipsilateral ear regardless of the hemisphere involved. In 11 cases involving temporal-lobe surgery, discrimination for filtered speech decreased for the ear contralateral to the side of the lesion even when Heschl's gyrus was not removed (Calearo and Antonelli, 1968).

Jerger (1960) reported findings that were in agreement with those of Bocca *et al.* (1954, 1955) in spite of having modified Bocca´s procedures. Jerger incorporated English PB word lists that were low-pass filtered at 500 Hz with a rejection rate of 17-dB/octave. In two reported cases of temporal-lobe pathology, Jerger (1960) noted that discrimination scores for the ear contralateral to the lesion were 30% poorer than those for the ear ipsilateral to the lesion. He (1964, cited in Jerger, 1973) further reported results for four groups of patients with CNS pathology. These results indicated (1) normal findings for the group with lesions restricted to the brainstem not involving the auditory system (N = 7); (2) a reduction in intelligibility of approximately 20% for the ear contralateral to the lesion in the group with brainstem lesions involving the CANS (N = 7), as well as the group afflicted with unilateral temporal-lobe lesions involving the auditory cortex (Heschl´s gyrus) (N = 6); and (3) for the group with cortical lesions not involving the auditory cortex, a reduction in discrimination scores, but to a lesser extent than for groups with lesions involving the CANS.

Lynn and Gilroy (1972) further modified Bocca's procedure by using Northwestern University (NU) Auditory Test #6 low-pass filtered at 500 Hz with a rejection rate of 34 dB in the first octave. Normal scores ranged from 40% to 70% with interaural differences of less than 8%. Findings for cases of temporal-lobe pathology provided further confirmation of the decrease in discrimination ability for the ear contralateral to the cortical lesion.

The low-pass filtered speech test used by Willeford in his CAP test battery and listed in Table A3 was developed by Ivey (1969). He used the Peterson and Lehiste (1962) words which were selected on the basis of previous research (Billingslea, 1963) which analyzed the intelligibility of low-pass filtered consonant-nucleus-consonant (CNC) words (500 Hz, 18-dB/octave rejection rate). While Billingslea (1963) reported that the ten CNC word lists resulted in homogeneous performance scores under filtered conditions, he cautioned that a marked difference existed in the intelligibility of individual items within each list. Ivey (1969) elected to set as a criterion for selection only those items which normal-hearing listeners could identify correctly 95% of the time under those filter conditions created by Billingslea (1963). The rationale for this selection procedure was that the task thus would be relatively easy for individuals without central-auditory dysfunction. Two 50-word lists were recorded on one channel of the tape with the carrier phrases monitored to peak at a 0-VU meter reading.

Willeford's Recommended Procedures

Two 50-word lists of specially selected monosyllabic (University of Michigan) CNC words (Ivey, 1969) - words selected so that they remain intelligible following electronic filtering with a cut-off frequency of 500-Hz and a rejection slope of 18-dB/octave - are presented at 50-dB SL (re: either the pure-tone average [PTA] or the spondee threshold [ST]). The presentation level of 50-dB SL (re: either the PTA or ST) was selected on the basis of performance-intensity (P/I) functions on normal-hearing adults, where it was noted that scores either peaked or asymptoted at this level (Gambino, 1963; Willeford, 1977). A difference in difficulty between List 1 and List 2 detected by Ivey (1969) and later confirmed by Willeford (1978), prompted the instigation of a correction factor of 10% to be added the scores obtained from List 2 to compensate for the difference in difficulty level between the two lists.

Table A3 contains the test items for the Low-Pass Filtered Speech test.

TABLE A3

LOW-PASS FILTERED SPEECH

List 1

1. home	11. jar	21. head	31. bite	41. coin
2. root	12. much	22. write	32. lot	42. lag
3. hide	13. kid	23. hire	33. dime	43. tire
4. more	14. war	24. gone	34. talk	44. cash
5. lap	15. have	25. dumb	35. coat	45. luck
6. phone	16. rain	26. book	36. shine	46. map
7. pole	17. curve	27. toad	37. bone	47. neck
8. mine	18. patch	28. choose	38. hot	48. watch
9. burn	19. moon	29. shock	39. search	49. fine
10. ride	20. car	30. such	40. lash	50. wash

List 2

1. wood	11. weak	21. wash	31. wish	41. should
2. hash	12. mire	22. vine	32. pan	42. loan
3. dab	13. loop	23. love	33. room	43. light
4. work	14. jet	24. bar	34. tone	44. wire
5. chum	15. what	25. juice	35. bug	45. sure
6. hush	16. chin	26. dock	36. tube	46. wet
7. hate	17. job	27. hole	37. bun	47. dish
8. which	18. turn	28. wheat	38. white	48. hair
9. joke	19. move	29. shade	39. pile	49. well
10. limb	20. word	30. neat	40. nose	50. pull

Findings Associated with Willeford's Test

Abnormally reduced scores (less than 74%) or between-ear difference scores greater than 10%, suggestive of cortical pathology have been reported in cases of peripheral hearing loss (Mittenberger, et al., 1978), learning disabilities (Willeford, 1974; Pinheiro, 1977; White, 1977) - including marked spatial problems (Musiek and Gerurkink, 1980) - and dyslexia (Welsh, et al., 1980), stuttering (Larimore and Willeford, 1980), and following decompression sickness (Mittenberger, et al., 1979). Studies on performance obtained on this task with learning disabled children indicated that left-ear performance tended to be slightly less than normal while right-ear performance was "borderline" normal (Pinheiro, 1977). Reduced left-ear performance also was apparent among a group of stutterers (Larimore and Willeford, 1980). In addition, of the four tests within the CAP battery, the LPFS test was the one most often failed in isolation (White, 1977).

4. CST

Review of Previous Research

The origins of the CST may be traced to early work by Broadbent (1954, 1962), Broadbent and Gregory (1964), Kimura (1961), Shankweiler (1966), and Shankweiler and Studdert-Kennedy (1967) which made use of various dichotically presented material including digits, musical notes, nonsense syllables, spondaically stressed words, synthetic and "natural" sentences. Reduced performance scores on dichotic tasks have been associated with cortical pathology (Kimura, 1961; Katz, 1962; Katz, *et al.*, 1963; Ivey, 1969; Jerger and Jerger, 1974; Keith, 1977; Lynn *et al.*, 1972; Willeford, 1977).

Early results, making use of dichotic digits, suggested a sensitivity of this test paradigm to cortical pathology (Kimura, cited in Brunt and Goetzinger, 1968). A laterality effect was first noted among normal-hearing listeners by Kimura, and was made evident with the right ear typically demonstrating superior performance scores over those achieved with the left ear. This right-ear advantage (REA) was evident in children as young as 6 years of age (Kimura, cited in Brunt and Goetzinger, 1968).

Berlin, et al. (1973), however, have indicated that the REA can be demonstrated reliably in children as young as 5 years of age.

The use of consonant-vowel (CV) syllables, in a dichotic presentation paradigm, has been advocated by other researchers (see for example Berlin and Lowe, 1972). A variety of interaural time-delays has been added to the strictly dichotic simultaneous onset condition, and these have been combined to form the basis of Berlin's dichotic nonsense syllable test (Berlin and Lowe, 1972). Interaural time-delays of 30, 60 and 90 ms have been used, and these have been purported to yield an REA to normal-hearing right-handed individuals, and also to be differentially sensitive to various cortical pathologies (Noffsinger and Kurdziel, 1979). The relative performance of the right ear compared to that of the left ear has been used to differentiate lesions of the left anterior, left posterior, and right temporal lobes, as well as lesions of the deep left parietal lobe, corpus callosum, anterior commisure and thalamus (Berlin and Lowe, 1972; Berlin, 1976; Berlin, 1981; Dermody and Noffsinger, 1976; Noffsinger and Kurdziel, 1979).

Further investigations by Tobey, et al. (1979), undertaken with learning disabled children, failed to demonstrate differences in performance (either in the magnitude of performance scores, the ear showing dominant performance, or the error pattern) between the disabled children and a control group of normal children. While identification of lagging stimuli was easier for both groups of children, as the stimuli's interaural time-delays became more disparate, only the normal children were able to improve correct identification of the leading stimuli. The learning disabled children, while able to maintain correct identification of the lagging stimuli, failed to show an improvement in their ability to correctly identify the leading stimuli, thus resulting in an apparent exaggerated lag effect. This pattern of performance by the learning disabled children was interpreted as being indicative of a reduced temporal processing ability, although the site of lesion that would result in such a deficit was not established.

One of the commonly used competing-message tests involves the use of spondaically stressed words, only portions of which are presented in a crude dichotic paradigm. This test has been identified as the Staggered Spondaic Word (SSW) test (Katz, 1962). The original ideal purpose of this test was to assist in the identification of central pathology even in the presence of concomitant peripheral impairment. The 40-item test consists of pairs of spondaically stressed words presented one each to the right and the left ear of the listener. The spondee pairs are time staggered at the ears such that one word leads the other by one syllable. Thus, the second syllable of the first word is presented approximately simultaneously with the first syllable of the second word. Complex scoring procedures associated with this test entail the comparison of each ear's monotic performance score, and each ear's crude-dichotic performance score over the 40 test items. The ear receiving the leading stimulus varied from item to item. Scoring patterns of the two ears are considered, corrected, and compared to normative data that are largely unpublished and which have originated virtually exclusively from Katz' laboratory. Attempts to obtain original raw data have met with failure. Ideally, it would appear that this test is sensitive to lesions of the auditory cortex. Differentiation of high- versus low-brainstem pathology also has been alleged.

The Synthetic Sentence Identification (SSI) test was developed by Speaks and Jerger (1965) and consists of third-order sentential approximations. The test since has been modified (Jerger, et al., 1968) to include competing messages. The purpose of this test was to provide the examiner with an instrument that would task the information handling capabilities of the auditory system in a more rigorous fashion than "natural" sentences. This added rigor was thought to be more likely to be sensitive to CANS lesions. Published research (Jerger and Jerger, 1975) indicates that this test, when used with a contralateral competing message, is sensitive to temporal-lobe lesions. However, when this test is used with an ipsilateral competing message, it appears to be sensitive to brainstem-level lesions.

The ubiquity of the competing test paradigm, and its apparent diagnostic usefulness, led Willeford to develop the CST subtest of his CAP test battery. Specifically, Willeford (1978) claims to have designed the test primarily for the assessment of the integrity of the CANS. The use of "natural" sentences was advocated in an attempt to simulate language constructions heard in everyday conditions, and the language level of the test items was purported to be adequate for usage with very young children. In addition, it was thought that individuals with low-level intelligence or with severely impaired literacy could cope with the task required for the valid administration of the test, although no data were found to support this contention. A number of researchers (Treisman, 1964; Berlin, et al., 1972; Speaks, 1975; Berlin and McNeill, 1976) have demonstrated that the presence of a competing signal whose content was similar to that of the primary signal increased the amount of competition thus making the task more difficult and, at the same time, more sensitive to CANS lesions. Because of these observations, Willeford constructed competing stimuli that were similar in length and content to those of the primary signals.

Willeford's Recommended Procedures

Two different sentences are presented, simultaneously, one to each ear with the primary sentence presented at 35-dB SL to one ear and the competing sentence at 50-dB SL (S/C=-15) to the other ear. The procedure then is reversed and a second list of sentences is presented so that the ear receiving the competing message in the first condition now is receiving the primary message.

Table A4 contains the test items for the Competing Sentences test.

TABLE A4

COMPETING SENTENCES

(a = primary signal,
b = competing message)

TEST No. 1

1. a. I think we'll have rain today.
   b. There was frost on the ground.

2. a. This watch keeps good time.
   b. I was late to work today.

3. a. I'm expecting a phone call.
   b. Please answer the doorbell.

4. a. The bus leaves in five minutes.
   b. It is four blocks to the library.

5. a. My mother is a good cook.
   b. Your brother is a tall boy.

6. a. Please pass the salt and pepper.
   b. The roast beef is very good.

7. a. There is a car behind us.
   b. This road is very slippery.

8. a. Leave the keys in the car.
   b. Fill the tank with gas.

9. a. It's always hot on the Fourth of July.
   b. Christmas will be here very soon.

10. a. We had to repair the car.
    b. You should really take a taxi.

TABLE A4 (continued)

TEST No. 2

1.	a.	The ice-cream sundae is very good.
	b.	We have chocolate and strawberry today.
2.	a.	Fasten your seat belt.
	b.	Get ready for take-off.
3.	a.	I think you need a band-aid.
	b.	You should see a doctor.
4.	a.	This is the latest style.
	b.	That fits you perfectly.
5.	a.	I will be back after lunch.
	b.	You may take this Saturday off.
6.	a.	I have seen this movie before.
	b.	This movie is not like the book.
7.	a.	Air-mail will get there faster.
	b.	Please answer on a postcard.
8.	a.	I think we have met before.
	b.	You probably don't remember me.
9.	a.	This train is going west.
	b.	All the cars are air-conditioned.
10.	a.	The children are playing baseball.
	b.	Football is an exciting game.

Findings Associated with Willeford's Test

Abnormally reduced performance scores (less than 90%) have been documented in patients afflicted with lesions of the posterior left temporal-lobe (Lynn and Gilroy, 1972), of the deep left parietal-lobe concomitant with corpus-callosum involvement (Lynn and Gilroy, 1972), and in some patients suffering from frontal-lobe lesions (Lynn and Gilroy, 1975). In addition, patients exhibiting learning disabilities (Willeford, 1976; Pinheiro, 1977; Musiek and Geurkink, 1980), and patients with vascular and degenerative disease (Lynn and Gilroy, 1975) have demonstrated abnormal and significant results. However, this test did not appear to be sensitive to lesions of the anterior temporal lobe (Lynn and Gilroy, 1972), nor to central-auditory pathologies associated with dyslexia (Welsh, et al., 1980).

APPENDIX B

CATEGORIZATION OF RECORDINGS

All the magnetic-tape recordings used during the present experiment were acquired from various clinics, and/or individuals within the past three years. The eight recordings acquired for measurement purposes were divided into two categories.

1. "OLD" -- Tape recordings were placed in this category if the length of time during which they had been used could not be ascertained, or if they were known to have been used more than five times.

2. "NEW" -- All tape recordings placed in this category either had never been used (N = 2) or had been used less than five times (N = 2).

APPENDIX C

PLAYBACK CALIBRATION OF THE TAPE RECORDER USED

Prior to any of the calibration measurements, the tape heads of the recorder/reproducer (Akai, Model 1722) were cleaned and degaussed in accordance with procedures specified by the manufacturer. A commercial head cleaner was used to clean the heads, while concentrated alcohol was used to clean the tape paths, capstan and roller.

SPEED CALIBRATION

The speed of the tape-drive unit was calibrated using a five-minute timing tape and stopwatch. Tape speed deviation of less than 0.3% for the preferred tape speed of 7.5 in/s was considered acceptable and is in accordance with the specifications promulgated by the National Association of Broadcasters (N.A.B., 1965). The range of the average speed of the recorder/reproducer was 7.52025 in/s to 7.47975 in/s. This error in the speed of the tape recorder will result in a maximum frequency deviation of 2.7 Hz at 1000 Hz. The specific speed deviations associated with the recorder/reproducer as determined prior to the measurement of each tape is contained in Tables C1 through C8.

PLAYBACK-LEVEL CALIBRATION

A block diagram of the instrumentation used for the calibration of the tape recorder/reproducer is contained in Figure C1. The output of the tape recorder was fed to the input of an impedance-matched loudspeaker. The output of the recorder also was monitored across the speaker with a microphone amplifier (Bruel & Kjaer, Type 2603).

Prior to the playback level calibration, the tape heads were aligned using a commercial reproduce/alignment calibration tape (Ampex, No. 01-31321-01), with 50- and 3180-us equalization. The tape heads were aligned for maximum output voltage using a 15000-Hz tone. The 700-Hz reference calibration tone recorded at playback level (10 dB below operating level) on the calibration tape was used to adjust the reproduce level of the recorder to obtain a 0-dB (relative) reading on the meter of the microphone amplifier. Bias and equalization adjustments were accomplished so that the output levels of the tape recorder at all other test frequencies would fall within the specifications. Results of this calibration are contained in Tables C1 through C8.

Figure C1

Block diagram of the instrumentation used for the calibration of the tape recorder and for the acquisition of data.

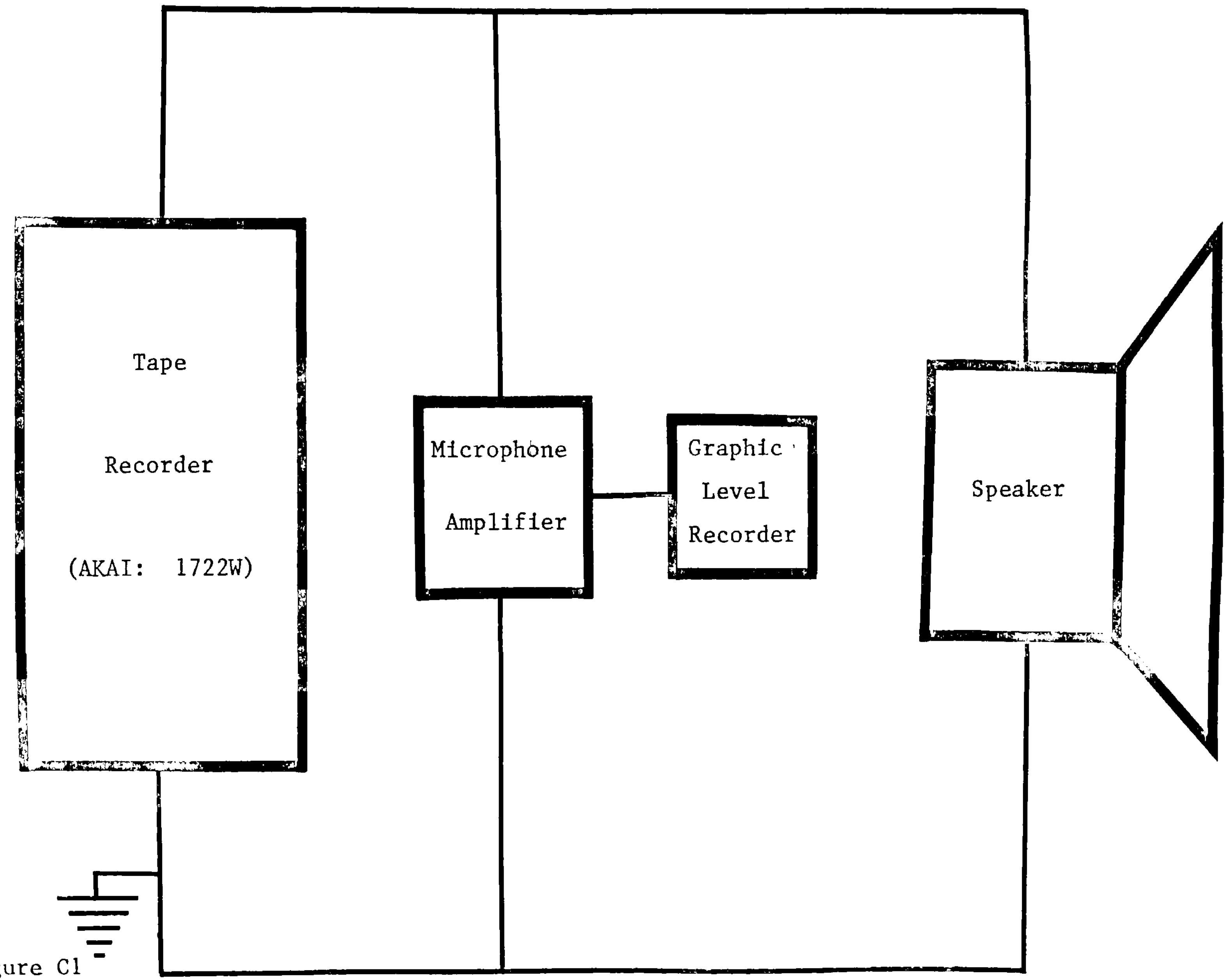

Tape
Recorder
(AKAI: 1722W)
Microphone
Amplifier
Graphic
Level
Recorder
Speaker

Figure C1

TABLE C1

Results of tape recorder playback calibration for Tape #1. (speed deviation = -0.27%)

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	+0.5	+0.2	+1 to -3
12000	+1.0	+0.5	+1 to -2
10000	+0.6	-1.0	±1
7500	+0.8	-0.1	±1
5000	+0.9	+0.4	±1
2500	+0.2	+0.2	±1
1000	-0.9	-0.8	±1
500	+0.1	0.0	±1
250	+0.7	+0.7	±1
100	+0.8	+0.3	±1
50	+0.7	+0.3	+1 to -3

TABLE C2

Results of tape recorder playback calibration for Tape #2. (speed deviation = +0.27%).

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	+0.5	+0.2	+1 to -3
12000	+1.0	+0.3	+1 to -2
10000	+0.4	-1.0	±1
7500	+0.8	0.0	±1
5000	+0.9	+0.3	±1
2500	-0.1	+0.1	±1
1000	-1.0	-0.6	±1
500	-0.5	+0.2	±1
250	+0.5	+0.9	±1
100	+0.4	+0.6	±1
50	+0.5	+0.3	+1 to -3

TABLE C3

Results of tape recorder playback calibration for Tape #3. (speed deviation = -0.17%).

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	0.0	-0.7	+1 to -3
12000	-0.1	-0.8	+1 to -2
10000	0.0	-0.9	±1
7500	+0.7	+0.1	±1
5000	+1.0	+0.6	±1
2500	+0.4	+0.8	±1
1000	-1.0	-0.2	±1
500	-0.4	+0.3	±1
250	+0.5	+0.9	±1
100	+0.5	+0.7	±1
50	+0.9	+0.8	+1 to -3

TABLE C4

Results of tape recorder playback calibration for Tape #4. (speed deviation = -0.03%).

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	-0.7	+0.1	+1 to -3
12000	-0.3	-0.1	+1 to -2
10000	-0.2	-0.5	±1
7500	+0.6	+0.3	±1
5000	+0.9	+1.0	±1
2500	+0.3	+0.9	±1
1000	-1.0	-0.2	±1
500	-0.5	+0.3	±1
250	+0.4	+1.0	±1
100	+0.9	+0.7	±1
50	+0.5	0.0	+1 to -3

TABLE C5

Results of tape recorder playback calibration for Tape #5. (speed deviation = -0.07%).

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	-0.4	-0.1	+1 to -3
12000	+0.5	+0.2	+1 to -2
10000	+0.3	-0.5	$\pm$1
7500	+0.6	+0.2	$\pm$1
5000	+0.9	+1.0	$\pm$1
2500	+0.2	+0.8	$\pm$1
1000	-1.0	-0.4	$\pm$1
500	-0.6	0.0	$\pm$1
250	+0.4	+0.7	$\pm$1
100	+0.5	+0.6	$\pm$1
50	+0.1	-0.3	+1 to -3

TABLE C6

Results of tape recorder playback calibration for Tape #6. (speed deviation = +0.13%).

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	-0.2	-0.3	+1 to -3
12000	-0.3	-0.2	+1 to -2
10000	+0.1	-1.0	±1
7500	-0.6	-0.1	±1
5000	+0.9	+0.8	±1
2500	+0.2	+0.3	±1
1000	-1.0	-0.5	±1
500	-0.7	0.0	±1
250	+0.8	+0.9	±1
100	+0.7	+0.7	±1
50	+0.4	0.0	+1 to -3

TABLE C7

Results of tape recorder playback calibration for Tape #7. (speed deviation = -0.03%).

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	-0.8	-1.0	+1 to -3
12000	-0.6	-0.9	+1 to -2
10000	-0.5	-0.9	±1
7500	+0.2	-0.3	±1
5000	+0.9	+0.4	±1
2500	+0.3	+0.2	±1
1000	-1.0	-0.7	±1
500	-0.3	0.0	±1
250	+0.7	+1.0	±1
100	+0.8	+0.6	±1
50	+0.4	-0.2	+1 to -3

TABLE C8

Results of tape recorder playback calibration for Tape #8. (speed deviation = -0.03%).

NOMINAL FREQUENCY (Hz)	CHANNEL 1 (relative dB)	CHANNEL 2 (relative dB)	TOLERANCE RE: NAB standards (relative dB)
700	0	0	--------
15000	-0.9	-1.0	+1 to -3
12000	+0.1	-0.9	+1 to -2
10000	0.0	-0.9	±1
7500	-0.4	-0.1	±1
5000	+1.0	+0.6	±1
2500	+0.8	+0.5	±1
1000	-0.7	-0.4	±1
500	+0.1	+0.2	±1
250	+0.9	+1.0	±1
100	+1.0	+0.7	±1
50	+0.5	-0.1	+1 to -3

APPENDIX D

DATA ACQUISITION

The instrumentation described in Appendix C was used to acquire the data for the present experiment. Additional details may be found in the 'PROCEDURES' section in the body of this document. The potentiometer range is 50 dB. The rectifier response was set to 'RMS' for all tests except the MASP in which case the response was adjusted to 'PEAK'. The writing speed was 200 mm/s and the paper speed was 30 mm/s. Using these instrument settings, hard-copy tracings were obtained of the recording levels of items on the tape recordings. As an example of the morphology of these tracings, some results for Tape #5 are contained in Figures D1 through D4.

Figure D1

Tracings for the DF, Items 6 through 8, List 1, Tape #5.

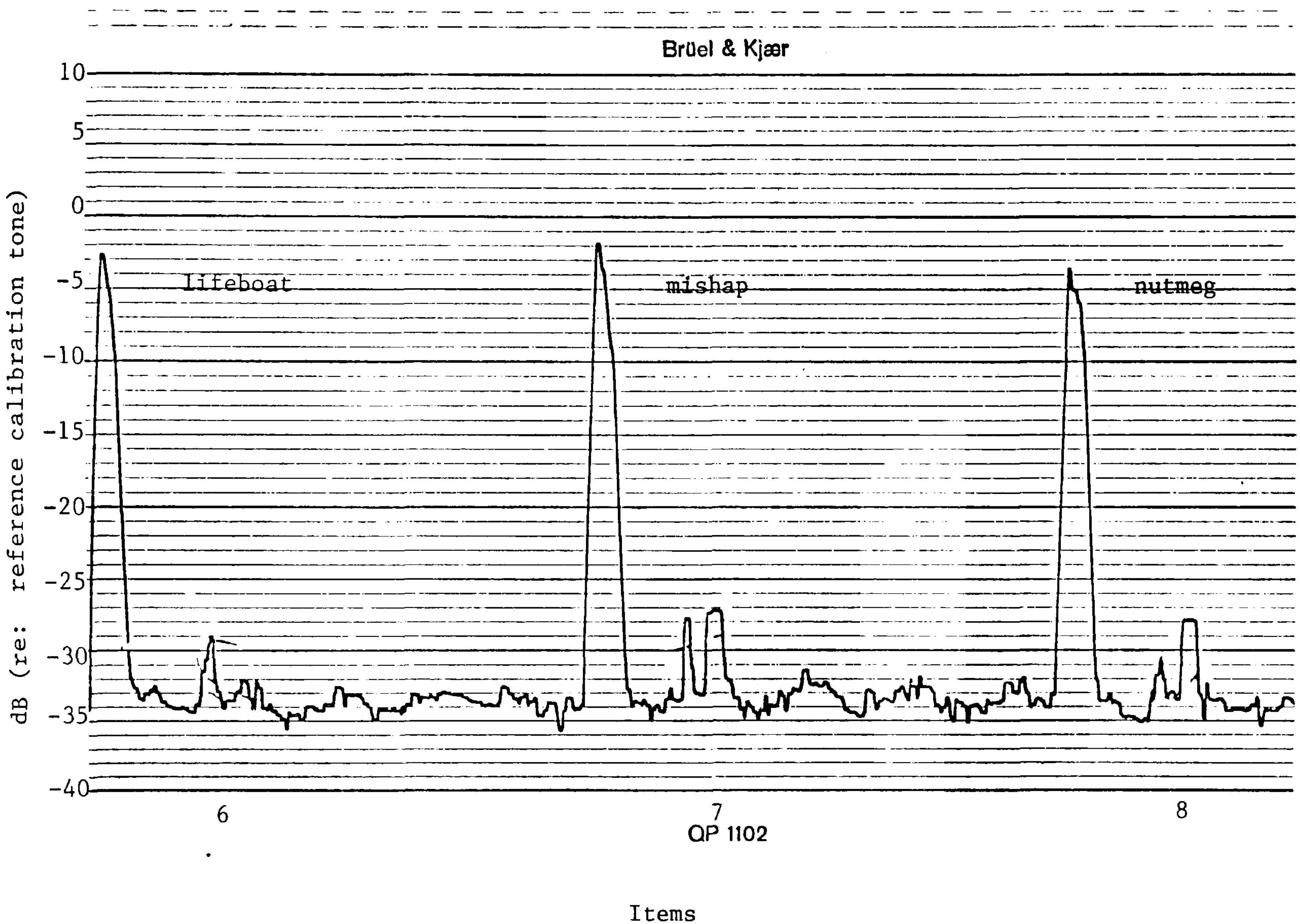


Figure D1 High band-pass filter.

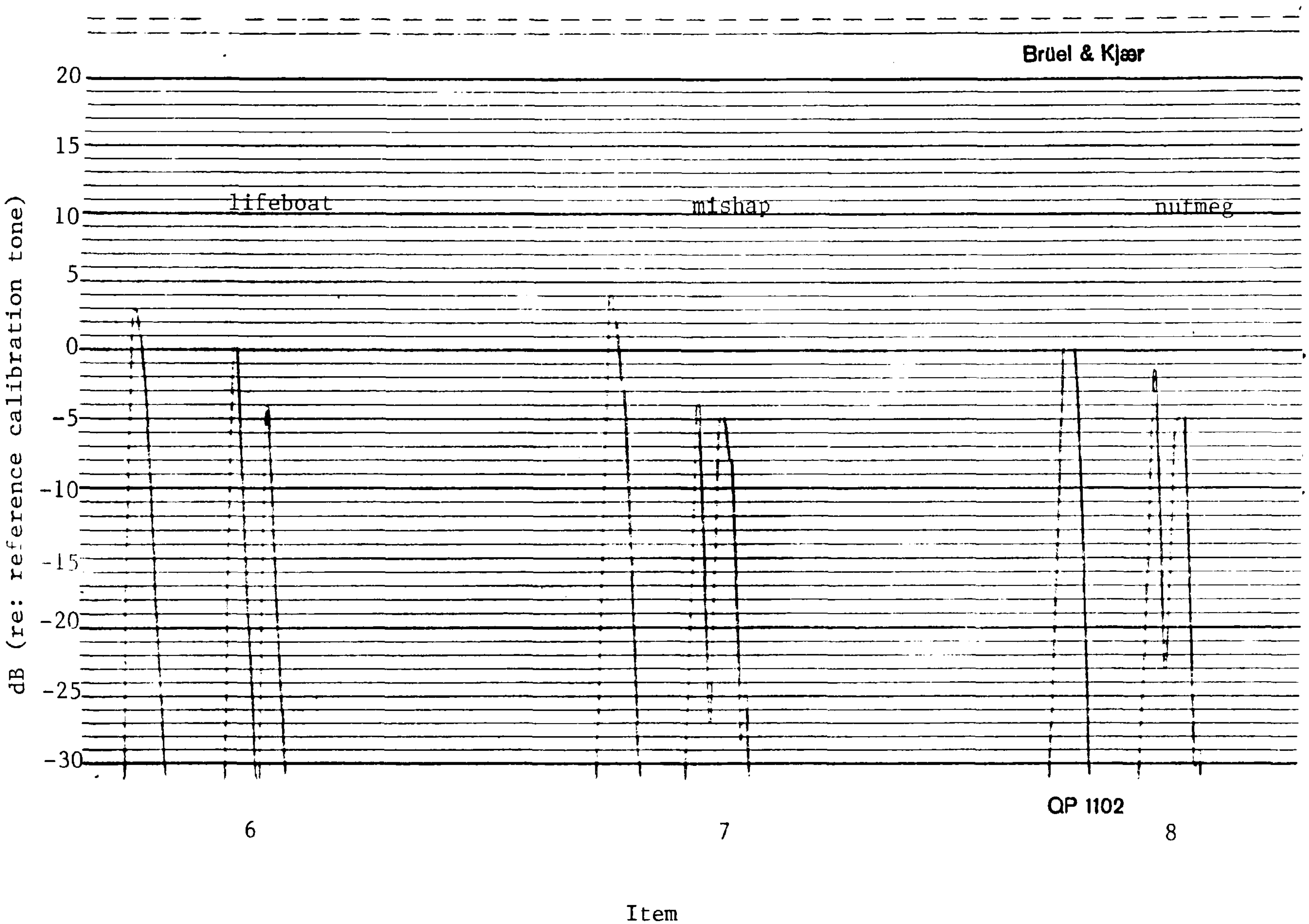
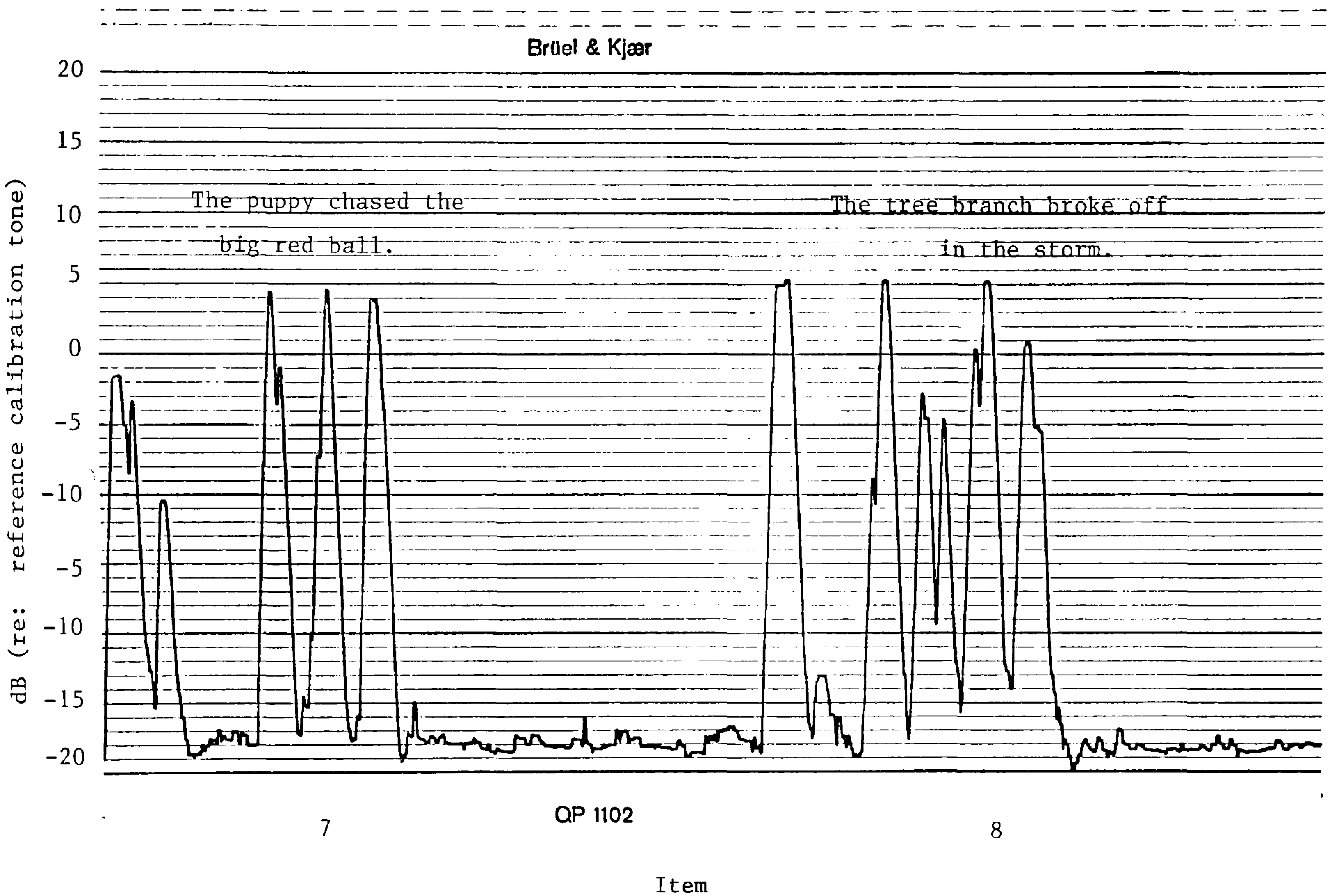


Figure D1 Low band-pass filter.

Figure D2

Tracings for the RASP, Channel 1, Items 7 through 8, List 1, Tape #5.


Brüel & Kjær
20
15
10
5
0
-5
-10
-5
-10
-15
-20
dB (re: reference calibration tone)
The puppy chased the
big red ball.
The tree branch broke off
in the storm.
QP 1102
7
8
Item

'igure D2

Figure D3

Tracings for the LPFS, Items 33 through 35, List 1, Tape #5.

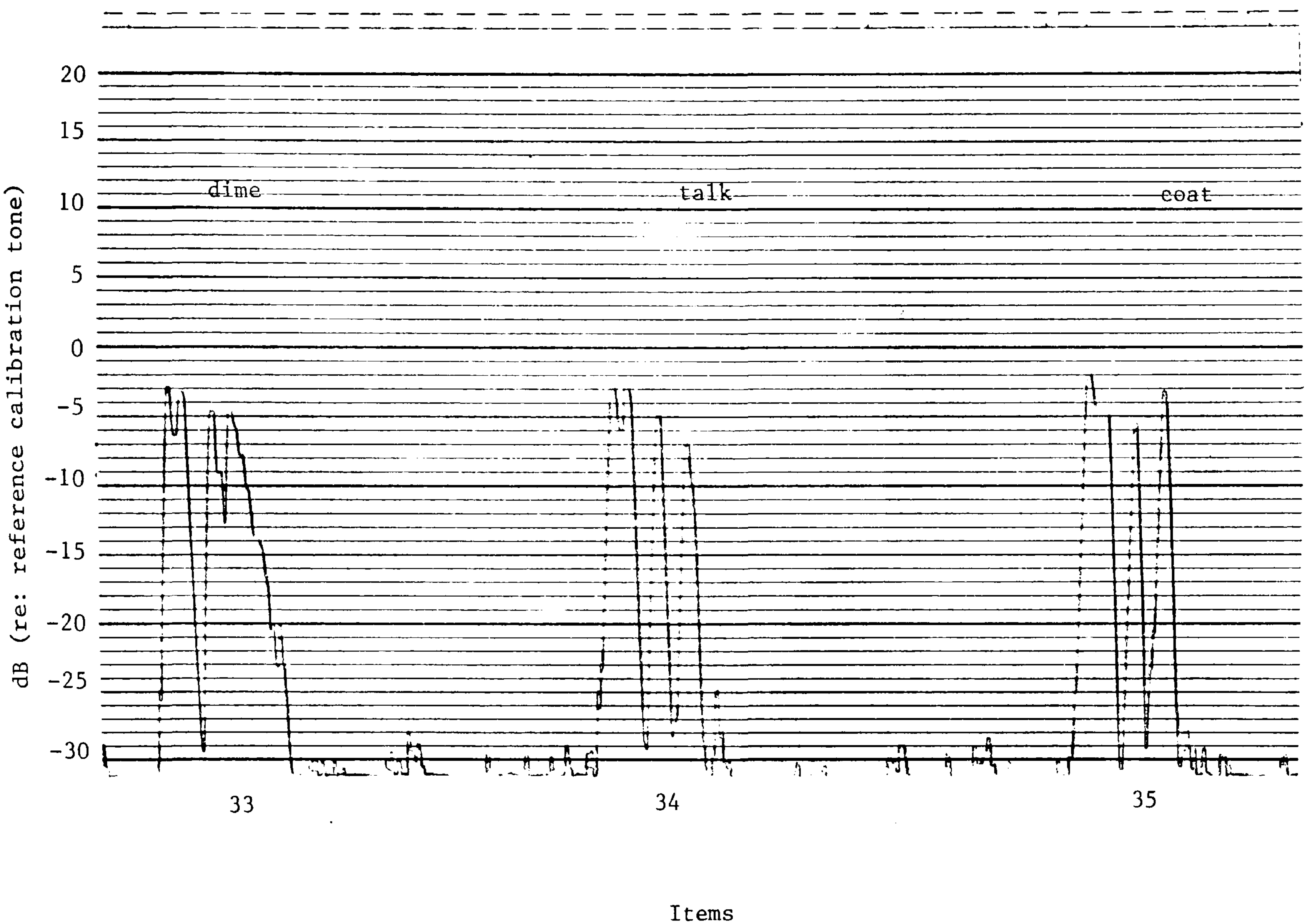

dB (re: reference calibration tone)
20
15
10
5
0
-5
-10
-15
-20
-25
-30
dime
talk
coat
33
34
35
Items

Figure D3

Figure D4

Tracings for the CST, Channel 1, Items 1 through 2, TEST 2, Tape #5.

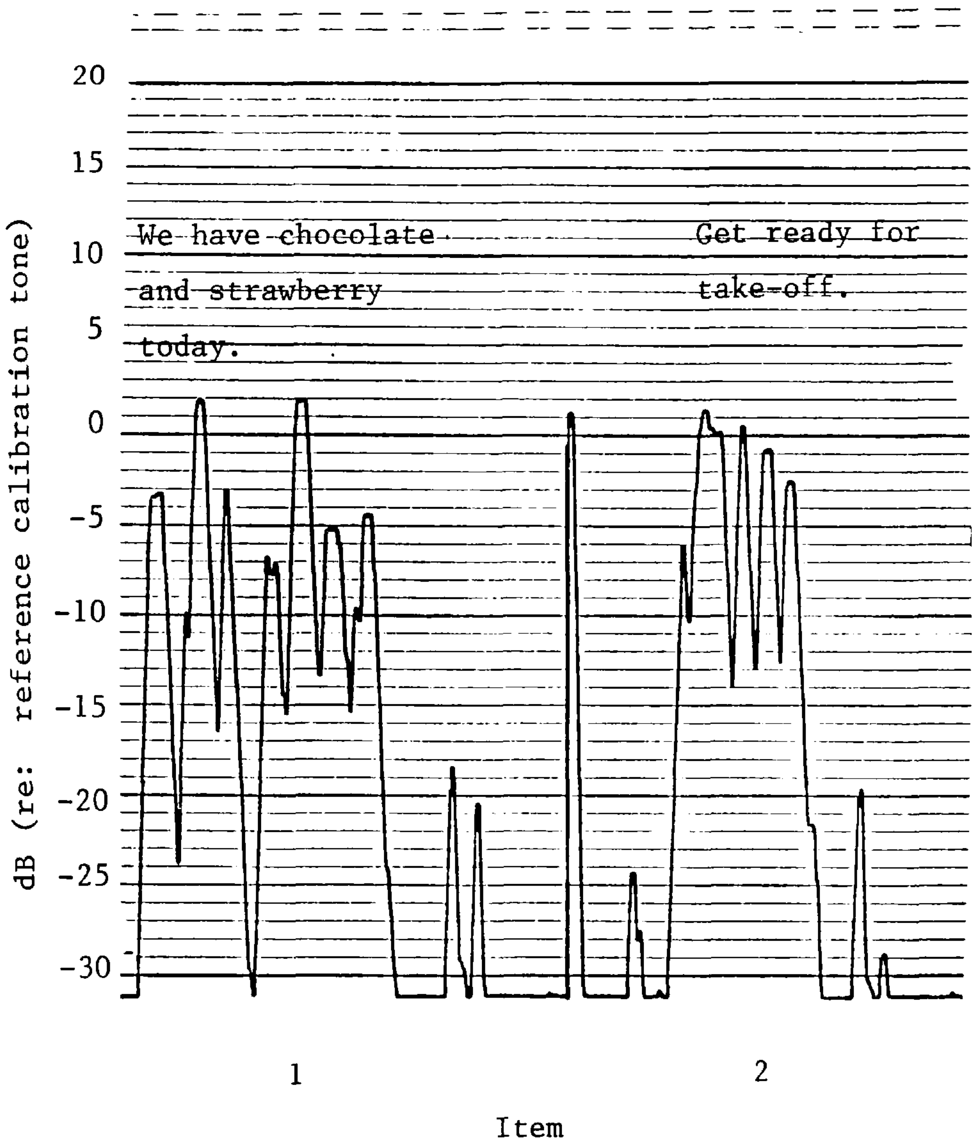

dB (re: reference calibration tone)
20
15
10
5
0
-5
-10
-15
-20
-25
-30
We have chocolate
and strawberry
today.
Get ready for
take-off.
1
2
Item

Figure D4

APPENDIX E

RAW DATA

Data are compiled by tape, and the tape number is indicated by a one-digit arabic numeral ranging from 1 to 8 as the first datum point of the first line of data of the tape. Tapes numbered 1 through 4 are considered "OLD", and tapes numbered 5 through 8 are considered "NEW" (see Appendix B). Each datum point represents the recorded level for the item of interest relative to the calibration tone level preceding the test. Each datum point is a real constant variable allocated a string of decimal digits immediately followed by a decimal point followed by a one-digit decimal fraction. Each datum point thus is represented by a five-digit number (including the decimal point as a digit).

The formats used for the several tests of the CAP test battery are as follows:

1. CST: Two lines of 15 data points each plus one line of 10 data points. The first line of data commences with a one-digit tape identification number. The first datum point

represents the value of channel 1, item 1, test 1. The second datum point represents the value of channel 2, item 1, test 1. The next two points are associated with values obtained for the second item of test 1 (channels 1 and 2 respectively). The last datum point is the value of channel 2, item 10, test 2.

2. LPFS: Six lines of 15 data points each are followed by one line of 10 data points. The first datum point represents the recorded level of item 1, list 1. The next point represents the level of item 2, list 1. The last datum point of the last line represents the value of item 50, list 2.

3. BF: Ten lines of 15 data points each are followed by one line of 10 data points. The first datum point represents the recording level of syllable 1, channel 1, item 1, list 1. The second datum point is associated with syllable 2, channel 1, item 1, list 1. The next two points are consistent with the first and second syllable respectively of channel 2, item 1, list 1. The next two data points are associated with syllables 1 and 2 of channel 1, item 2, list 1. The last datum point is the recording level of syllable 2, channel 2, item 20, list 2.

4. RASP: The data are contained in four lines of 15 data points each. The first datum point represents the value associated with channel 1, item 1, list 1. The next represents channel 2, item 1, list 1. The third datum point represents channel 1, item 2, list 1. The last datum point represents channel 2, item 10, list 3.

The data thus formated are shown in Table E1.

TABLE E1

RAW DATA

```
1-01.5-02.0-02.5-02.0-01.5-01.5-03.5-00.5-01.5000.0-01.5-01.0-02.5-01.5-01.5
-00.5-02.5-02.0-03.0-01.0-02.5-01.5-02.0-01.0-02.5-02.5-02.0-03.0-01.5-01.5
-02.0-01.5-00.5-01.0-12.0-02.5-05.5-01.5-05.0-02.0
-09.0-08.5-12.0-07.5-10.0-10.5-08.5-08.0-08.0-08.0-11.0-08.0-11.0-08.5-12.0
-09.0-08.0-12.5-06.0-11.5-11.0-08.5-13.0-08.5-08.5-09.0-10.5-08.0-12.5-12.5
-11.5-10.0-10.5-12.0-09.5-12.0-09.0-12.0-10.5-10.0-10.5-11.0-12.5-12.5-08.5
-09.0-09.5-10.0-12.0-09.0-10.5-12.0-12.0-09.5-11.0-12.5-12.0-10.0-11.5-09.0
-10.0-10.0-09.5-12.0-10.0-09.5-10.5-11.0-09.0-10.5-10.0-11.5-09.5-13.5-09.5
-12.0-09.5-09.5-08.5-08.5-08.5-09.0-09.0-10.0-10.5-08.5-09.5-10.5-12.5-09.5
-09.0-09.0-10.5-12.0-10.5-11.5-08.5-11.5-11.5-09.5
-31.0-34.0-06.5-13.5-37.5-34.0-11.0-09.5-33.0-37.0-11.5-10.5-36.0-36.0-08.5
-08.5-32.5-34.5-12.5-10.0-33.0-33.5-10.0-10.0-30.5-33.0-12.5-11.0-34.0-36.0
-08.0-13.5-35.0-36.5-10.5-10.5-34.5-35.5-09.5-09.5-34.5-32.0-07.5-12.5-34.5
-35.0-24.0-12.0-34.0-35.5-07.0-14.0-31.0-33.5-08.5-09.0-34.0-37.0-23.5-07.0
-37.5-32.5-15.5-11.5-33.0-33.0-09.5-14.0-37.5-34.0-09.0-06.5-38.0-39.0-07.0
-08.5-33.0-36.5-07.0-08.5-37.5-31.0-09.0-13.5-37.5-38.0-12.0-12.5-28.0-26.5
-08.5-20.5-32.0-34.5-10.5-11.0-36.0-36.5-07.0-11.5-33.0-36.5-06.5-10.5-33.5
-29.5-13.0-10.5-37.5-30.5-10.0-10.0-33.5-37.0-08.0-13.0-34.5-37.0-07.5-05.5
-34.0-37.5-15.5-09.5-34.0-37.5-07.5-13.0-32.5-31.0-08.5-13.5-29.5-34.0-06.5
-08.5-37.5-34.0-07.0-07.0-32.0-36.5-07.0-14.0-34.0-32.0-05.5-11.0-36.5-30.0
-12.5-07.5-31.0-35.5-06.5-15.0-36.5-34.0-08.0-08.0
-02.0-04.5000.5-05.0-02.0-05.5-03.0-04.0-03.5-04.5-04.0-05.0-03.0-06.5-04.0
-05.5-03.0-12.5-04.0-18.0-03.5-02.5-05.0-03.0-01.5-05.0-03.0-05.0-01.5-04.5
-03.5-04.5-03.5-03.5-04.0-04.5-02.5-05.5-03.5-05.5-03.5-05.0-04.5-03.0-03.0
-04.5-02.5-04.5-02.0-06.0-03.5-05.5-04.0-05.5-02.0-02.5-05.5-03.0-02.5-04.5
```

TABLE E1 (continued)

```
2000.0-00.5000.0-02.0002.0001.0000.5-00.5001.5001.0001.0-03.0000.5-01.5-01.0
000.5000.0001.0-02.0000.5-01.5001.5001.0000.5001.0001.5000.5-01.5003.0002.0
003.0001.5-00.5003.0002.5001.0003.5004.0002.0002.5
-01.5000.0-06.0-01.0-04.0-03.5-01.5-02.0-00.5-02.0-02.0-02.5-02.5-01.0-05.5
-02.0-00.5-02.5001.0001.5-03.0-00.5-06.5000.0-01.0-00.5-03.5000.0-05.0-06.0
-04.0-04.0-03.0-05.0-01.0-05.0-01.0-06.0-02.0-03.5-02.0-04.5-05.5-04.0-02.0
-03.0-03.5-03.5-04.5-02.5-02.5-07.0-04.0-01.5-03.0-05.0-05.5-02.5-01.0-02.0
-02.0-02.0-01.0-03.5-02.0-01.5-01.0-04.0-02.5-04.0-03.0-04.0-02.0-05.5-01.5
-03.0-02.0-01.5-01.5-02.0-02.0-02.0-02.0-03.0-02.0-02.0-02.5-04.0-05.0-03.0
-01.0-03.0-03.5-05.0-02.0-05.5000.0-04.5-03.0000.0
-27.0-32.0001.5-04.5-32.5-28.0-01.0000.5-28.0-35.0-06.0-03.0-31.5-32.5000.0
-09.0-27.0-30.5-04.5-01.0-29.0-31.0-02.0-02.5-27.0-35.0-02.5-02.5-30.0-33.5
000.0-05.5-30.0-34.0-04.5-02.5-30.0-33.0-05.0-08.5-30.5-28.5-00.5-04.5-32.0
-32.0004.0002.0-30.0-34.0000.0-06.0-28.0-30.0001.0-02.0-30.0-33.5-13.5000.5
-32.0-30.0-06.0-01.5-30.0-30.0-01.0-05.0-34.0-31.5000.0002.5-35.0-35.0001.0
-01.5-28.5-34.0002.0-00.5-31.0-29.0-02.0-06.0-33.0-35.0-08.0-06.0-23.0-27.0
000.0-12.0-27.5-28.5-01.0-02.0-33.0-33.0003.0-02.0-33.5-28.5002.5-02.0-28.5
-27.0-01.0-03.0-32.5-27.0001.5-02.5-28.5-33.5-01.0-03.0-29.5-31.5000.5-08.5
-28.5-34.5-02.0-00.5-32.0-31.5000.5-05.0-28.0-28.0-01.0-03.5-25.0-33.0002.0
-01.0-32.5-31.5001.5-10.0-31.0-29.0002.0-05.0-29.5-30.0-04.0-04.0-30.0-26.0
-02.0000.0-26.0-32.0002.0-06.0-33.0-31.5000.0000.0
007.0011.0010.0004.0005.5002.0008.5009.0002.0005.0-00.5008.5005.0010.0005.0
008.0003.0005.0007.5006.0011.0003.0007.0005.0005.5007.5007.5007.5008.0006.5
009.0006.5005.5007.0007.0005.5005.0007.0-01.5010.0006.5004.5009.0007.0004.5
007.5005.5009.0003.0009.5006.0006.5007.0009.5007.5009.0009.5006.5006.0003.0
```

TABLE E1 (continued)

```
3014.5012.0014.0011.0014.5013.0013.0013.0015.0014.0014.5011.5013.5012.0014.5
014.0013.0013.0014.0013.5010.5014.0012.0013.0013.0014.0015.5013.0014.0013.5
014.0014.5015.0013.0002.5013.0007.5012.0011.0007.0
001.5002.0-05.0000.5-04.0-04.5-01.5-01.5-00.5-02.0-03.0-01.5-05.0-02.0-04.5
-01.0-01.5-04.0002.5-03.5-02.0-01.5-06.0000.0000.0-01.5-04.0000.0-06.0-07.5
-04.5-03.5-02.5-05.5-03.0-05.5-01.0-05.5-04.5-04.0-02.5-04.0-05.5-04.5-02.0
-02.0-02.0-04.0-05.0-02.0-03.5-07.0-05.5-03.0-05.0-05.5-06.0-03.0-04.0-02.0
000.0-04.0004.0000.5002.5002.0002.5002.5003.0-01.0001.5001.0002.5-01.0003.5
001.5004.0004.0004.0004.0003.5003.5004.0003.0003.0004.0003.0001.0000.5000.0
003.5004.0002.0000.0003.0000.0006.0000.5002.5002.5
-25.5-26.5004.0-03.0-27.0-27.5-02.0000.0-28.0-25.5-07.0-02.5-28.0-27.5-01.0
-09.0-24.5-23.5-05.5-01.0-28.0-29.0000.0000.0-28.0-27.5-02.5-02.5-29.0-28.5
-02.0-04.0-27.0-28.0-04.5-03.0-27.0-28.5-07.0-07.0-27.0-26.5-00.5-03.0-28.0
-30.0003.0001.0-27.0-27.5-01.5-03.0-26.5-28.0003.0001.0-27.5-26.0-11.0001.0
-27.0-27.0-05.0-02.0-28.0-28.5-01.0-05.5-28.0-27.0003.0004.5-28.5-28.5003.0
001.0-27.0-27.5005.0000.0-27.5-27.5001.0-05.0-29.0-28.0-07.5-03.5-25.0-27.0
-01.5-10.0-26.0-27.0002.0-00.5-28.0-29.0004.0-01.0-26.0-27.0004.0-01.0-25.0
-26.5000.0-01.5-27.0-25.5001.5-01.0-26.5-27.5002.0-03.5-27.5-27.5004.0-07.0
-25.0-26.5000.0002.5-27.5-27.5003.0-04.0-25.5-25.5000.5-03.0-26.5-27.5002.5
002.0-28.0-28.0003.5-08.0-26.0-28.0004.0-04.0-28.0-26.0000.5-03.0-25.0-26.5
-01.0001.5-25.5-26.5003.5-06.5-28.0-27.0002.5003.0
007.0014.0012.0002.0009.5016.0011.0010.5012.0013.5008.0010.5007.0010.5009.0
007.5008.0011.0004.5008.5006.5011.0011.0007.0010.0010.0014.0009.5009.5010.5
011.0009.0008.0010.5010.0007.0008.5010.0013.5009.0010.5012.0012.5005.0007.0
008.0011.0007.5008.5009.0008.5011.0010.5007.5010.0013.5010.0012.0010.0011.0
```

TABLE E1 (continued)

4009.0005.5006.0004.5004.5007.5003.0006.0005.0008.5004.5005.0004.0002.5002.5
008.5005.5005.5004.5005.5004.0006.5006.5007.0007.5005.5006.5007.5007.0007.5
007.5007.5006.5005.5004.5007.5-01.0008.0-00.5004.0
-02.5-01.5-07.5-01.5003.5001.5-00.5-02.5000.5-03.5001.5-03.5002.5-02.0-06.0
-03.0-04.5001.0000.0-01.0-02.0-03.0-06.0-02.5-00.5-01.0001.0001.5-01.5001.5
-04.0-05.5-04.5001.0001.5000.5-02.5-05.0002.5-01.5000.5-06.0-05.0000.5-03.0
-02.0-04.5-05.0000.0-04.0-04.0-08.0-06.0-03.5002.5-07.5-06.0-04.0000.0-03.0
-03.5-03.5-04.0002.0-03.5002.0001.0001.5-03.5-06.0-04.5-04.0-03.0-03.5003.0
-05.0-01.0-02.5001.5-03.0-03.5-03.5-04.0000.5-04.0001.0-03.5-05.5-05.0-04.0
002.5-03.5-04.0-06.0001.0-07.0002.0-03.5-05.0015.0
-28.0-30.0018.5-04.0-29.0-27.5-01.0000.0-27.0-29.5-03.5-01.5-31.0-28.5018.5
018.5-27.0-29.0-04.0-00.5-27.0-30.5011.0011.0-26.0-30.0-07.0-01.0-29.0-30.5
018.0-04.0-31.0-31.0014.0-01.0-27.5-30.0018.5018.5-29.5-27.5019.5-04.5-30.0
-30.0-12.0-03.5-31.5-33.0002.0002.0-30.0-32.0009.5-01.0-31.0-34.0-08.0010.0
-27.5-30.0-03.0-01.0-27.0-27.5012.0-05.0-28.5-28.5-03.0002.0-31.0-32.0000.5
001.5-31.5-31.5001.0-01.5-27.5-31.5-01.0-07.5-30.0-30.5-02.0-05.5-28.0-32.0
001.0-01.0-24.0-26.0-01.0-04.0-30.0-30.0003.0001.0-26.5-29.5002.0001.0-20.5
-27.5-01.5-05.0-25.0-27.5001.0-02.5-29.5-32.0000.0000.5-23.0-29.0001.0-06.5
-24.0-29.0-03.0-01.0-20.0-31.5009.5-06.0-15.5-25.5-01.0002.5-30.5-30.0002.5
000.0-27.0-30.0000.0-10.0-22.5-29.0000.0001.0-26.5-31.5-05.5000.0-25.5-29.5
-02.0002.0-29.5-29.5015.0000.0-30.0-29.0016.5-01.0
003.0006.5008.5001.0002.0003.5005.0006.0001.5002.0001.0003.0001.0003.0003.0
006.5003.5002.5000.5001.5004.0003.0004.5005.0004.0009.0003.5005.0004.5004.5
005.5004.0003.0004.0007.0003.5004.0005.5-02.0004.5006.0001.0006.5004.0007.0
003.0005.5001.5004.0004.0006.0001.5004.5004.5006.0006.5004.0005.5003.5002.5

TABLE E1 (continued)

```
5001.0-03.0-03.0-10.0-00.5-02.5-04.0-03.5-02.0-04.0-03.0-06.5-02.0-07.5-04.0
-01.5-02.0-03.5-02.5-04.0-05.0-00.5-01.5-02.0-02.0-03.5-01.0-05.5-02.0-02.5
-02.0-01.0-01.5-01.0000.0-01.5002.0-03.5000.0000.0
-03.5-01.5-08.0-02.5-06.5-06.0-02.5-04.0-02.0-03.5-03.5-04.5-05.0-02.5-07.0
-04.0-02.5-05.0-01.0-05.0-05.0-02.0-08.5-02.0-02.5-02.5-06.0-02.5-07.5-08.5
-06.0-05.5-04.5-07.0-03.0-07.5-03.0-07.5-05.0-06.0-04.5-06.5-07.5-06.0-04.0
-04.5-05.5-05.0-07.0-04.0-04.5-09.0-06.0-03.5-05.0-07.0-07.5-04.5-03.0-04.0
-04.0-10.0-03.0-06.0-04.0-04.0-04.0-06.0-05.0-06.5-05.0-06.5-05.0-08.0-04.0
-05.5-04.0-04.0-04.0-04.0-03.5-05.0-04.0-04.5-04.5-04.0-05.0-06.0-07.5-07.5
-04.0-05.0-05.0-07.5-04.0-07.5-02.5-06.0-05.0-03.0
-27.5-31.5000.0-06.0-30.0-27.5-04.0-02.5-27.5-22.5-08.0-05.0-30.0-31.0-02.5
-10.5-27.0-30.0-07.0-03.0-29.5-31.0-04.5-04.5-27.0-33.0-05.0-04.5-29.0-32.0
-03.0-07.5-31.0-30.0-06.0-04.5-30.0-31.0-08.0-10.0-29.0-29.0-03.5-06.5-32.0
-30.5-13.5-04.5-30.5-33.0-03.0-07.5-27.5-29.0-00.5-02.5-29.0-32.5-15.0-02.0
-31.0-29.0-08.0-03.5-30.0-30.0-03.0-07.5-32.5-30.5-01.5000.5-34.0-33.5-01.0
-03.0-28.0-32.5000.5-03.0-31.0-28.0-03.5-07.5-31.5-32.5-11.0-07.5-23.0-32.0
-02.5-14.0-27.0-28.5-02.0-03.5-31.5-31.5001.0-04.0-29.0-32.0000.5-04.0-27.5
-27.0-03.5-04.5-30.5-28.0-01.5-04.5-31.0-32.5-02.5-06.0-30.5-31.5-00.5-10.5
-29.5-32.5-03.5-02.0-31.5-33.5-01.5-06.5-29.5-29.0-03.0-05.0-26.5-32.5-00.5
-02.5-32.0-32.5-00.5-11.5-29.0-34.0-00.5-07.0-30.0-29.5-06.0-06.5-31.5-27.5
-04.5-02.0-28.0-31.0000.0-08.5-31.5-30.5-02.0-01.0
001.5005.0004.0-01.5003.5002.5002.5001.5-01.0006.0001.0003.0003.0004.0-00.5
002.5004.0001.5003.0003.0002.5005.0002.5001.5003.5005.0005.0004.5005.0002.5
004.5004.0002.0002.5003.0001.5001.5002.0-04.0002.5-00.5003.0006.0004.0002.5
005.0003.5000.5000.0004.5002.0005.0005.0003.0004.0005.0002.0004.0001.5000.0
```

TABLE E1 (continued)

```
6000.5000.0000.5-03.0000.5-02.0000.5002.0001.5-01.5000.5-01.0001.0-04.5000.0
-01.0000.5-02.0000.5-00.5002.0000.0000.0000.0000.0-02.0001.5-02.0001.0-00.5
000.0-01.0000.0-00.5-08.0-00.5-07.0-02.5-04.0-01.5
-09.0-07.0-14.5-08.5-12.5-12.0-07.0-12.0-07.0-07.0-10.0-09.5-10.5-05.5-14.0
-09.0-07.0-11.5-08.5-10.0-11.0-05.5-15.0-08.0-09.0-06.5-11.0-08.0-11.5-11.5
-10.0-10.5-11.0-12.0-08.0-14.0-08.5-14.0-08.5-11.0-10.0-13.0-13.0-10.5-10.0
-13.0-12.5-10.0-13.0-10.0-09.5-14.0-12.0-09.5-13.0-13.5-13.0-09.5-08.0-10.0
-11.0-12.0-08.5-11.0-08.5-11.0-09.0-13.5-12.0-12.5-10.5-15.0-09.0-15.0-10.5
-12.0-08.5-09.0-10.0-12.0-09.5-12.0-11.5-13.0-11.0-10.0-11.5-13.0-14.0-12.0
-09.0-11.5-13.0-11.5-10.0-13.5-08.0-12.0-10.5-07.0
-28.0-31.5-04.0-07.0-31.0-28.0-06.5-03.5-27.5-32.0-10.0-06.5-30.0-30.0-02.5
-12.0-28.0-29.0-09.5-05.5-29.5-31.0-07.0-07.5-27.0-32.0-06.0-08.0-29.0-30.5
-03.5-10.0-30.0-31.5-08.0-07.0-29.5-31.0-07.5-13.0-29.5-28.5-03.5-10.5-30.0
-30.5-16.5-06.5-29.5-32.0-03.0-11.0-27.5-29.5-04.5-06.5-29.0-30.5-19.0-03.0
-31.5-30.0-10.5-06.0-30.0-30.0-05.5-11.0-31.0-31.0-03.5-01.0-31.5-31.5-04.5
-07.0-27.5-31.0-02.5-05.5-31.5-28.5-06.0-10.0-31.0-32.0-10.5-10.5-25.0-32.0
-05.0-18.0-28.5-30.0-07.0-08.0-31.0-31.5-01.5-08.0-28.5-28.5-02.5-06.5-29.0
-28.0-06.5-06.5-31.0-28.0-03.5-07.5-29.0-32.0-04.5-07.5-29.5-26.0-03.5-13.5
-28.5-31.0-07.0-04.5-30.5-31.5-04.0-08.5-29.5-28.5-05.0-08.0-26.0-32.0-02.0
-06.0-30.5-30.5-03.0-16.0-30.0-32.0-03.5-12.0-30.0-29.0-10.0-08.0-30.0-28.5
-06.0-03.5-29.5-27.0-02.5-10.0-30.5-29.0-04.5-04.0
-03.0003.5-03.0-05.0-01.5-05.5000.0-04.0-02.0000.0-03.0-06.5-03.0-06.5-02.0
-02.0-02.5-05.5-07.0-03.0-04.0-04.5000.0-01.5-01.0-04.0-00.5-05.0-02.0000.0
000.0-05.5001.0-05.5000.0-02.0-05.0-05.5-06.5-04.5-03.5-00.5-04.0-01.5-01.0
-02.5001.5-03.5-06.0-02.0000.0002.0001.0-02.5000.5-00.5-02.5-04.0-04.0-07.0
```

TABLE E1 (continued)

```
7000.5-02.5-01.5-07.0000.5-01.5-01.0-01.5000.5-02.0-01.0-05.0001.0-03.0-01.5
-00.5-02.5001.0000.5-01.0-01.5-02.5000.0000.0000.0-01.5001.5-07.5-00.5-01.0
-01.0-00.5-01.0-01.5-09.5-01.0-04.5-03.0-02.0-03.0
-11.0-08.5-15.5-10.0-13.5-13.0-10.0-12.0-09.5-11.0-11.5-12.0-12.5-10.0-15.0
-11.5-10.0-13.0-09.0-13.0-12.5-09.5-16.0-09.5-10.0-10.0-13.5-10.0-14.5-15.0
-13.5-13.0-12.0-14.5-10.5-15.0-10.0-15.5-12.0-13.0-11.5-14.5-15.5-14.0-11.5
-13.0-13.5-12.5-15.0-11.5-13.5-17.0-14.0-11.0-13.0-16.0-16.5-13.0-12.0-11.5
-12.0-13.0-11.5-13.0-11.5-12.0-11.0-14.0-12.5-14.5-13.5-14.5-11.5-16.0-12.0
-13.5-11.5-12.0-12.5-11.5-11.0-12.5-12.0-13.5-12.5-12.0-12.5-13.5-15.0-13.5
-10.5-13.0-14.5-15.0-12.5-15.0-10.0-14.0-13.5-10.0
-25.0-30.0-08.5-13.5-30.0-25.0-10.0-08.5-25.5-33.0-13.0-11.0-29.5-30.5-08.5
-18.5-25.0-26.0-12.5-11.5-27.5-30.0-10.5-12.5-23.0-31.5-11.0-12.0-28.5-30.0
-07.0-15.0-30.0-31.0-12.5-11.0-27.5-28.5-11.0-15.5-28.5-25.5-07.5-15.0-28.0
-29.0-25.0-12.5-27.5-32.0-06.5-15.5-25.5-28.0-08.5-11.0-26.0-31.5-23.5-07.5
-30.5-27.5-15.0-11.0-28.5-27.0-10.5-14.5-31.0-29.5-09.0-06.0-31.5-32.0-08.5
-12.5-26.0-29.0-07.5-10.5-29.5-26.5-10.0-16.5-30.5-30.5-14.0-15.0-19.5-23.5
-06.5-21.5-27.5-28.5-12.5-12.5-31.0-30.0-05.5-12.5-27.0-33.0-05.5-11.0-26.0
-25.5-11.0-11.5-30.5-26.0-08.5-12.0-28.5-32.0-10.0-12.5-28.5-31.0-09.0-17.5
-25.0-31.5-12.5-09.5-29.0-30.0-09.0-14.5-26.5-25.0-10.5-13.0-23.0-31.0-06.0
-10.5-30.5-28.5-07.0-20.0-27.0-32.5-07.5-15.5-27.5-27.0-13.5-13.0-30.0-25.0
-10.0-08.5-26.0-30.0-06.5-14.0-30.0-23.0-09.5-08.5
003.5008.0003.5002.0005.0005.0003.0004.5000.0007.5002.0003.0003.0004.0002.0
003.0-06.0003.5-08.0006.0004.0006.5006.0002.0005.0005.5008.0005.0004.5009.0
005.0004.0004.0003.0005.0003.0003.0005.0-00.5003.5003.5007.5006.0004.0003.0
004.0004.0005.0003.0006.5005.0006.0005.0005.0006.0005.5003.5008.0005.0004.0
```

TABLE E1 (continued)

```
8008.0004.5004.0003.5003.5005.0002.5006.5004.5008.0005.5004.0005.5003.0003.5
009.0005.0003.0005.0005.0003.5006.0006.0006.5005.5004.5005.0007.0007.0006.5
008.0006.5006.5006.0004.5006.5001.0006.0002.5007.0
-02.0-00.5-07.0-00.5-05.5002.5001.0-01.5001.5-03.0002.5-02.5004.0003.5-05.5
-02.0-02.5001.5001.0-00.5-01.5-02.0-04.5-01.5000.0000.0002.0002.5000.5002.0
-03.0-04.5-04.0003.0002.5002.0-01.5-04.5003.0004.0002.0-01.5-04.0002.0-02.0
-02.5-03.5-04.0001.0-02.5-03.0-05.0-05.0-02.5002.5-06.0-06.0-03.0001.5-03.5
-03.0-05.0-02.0002.5-03.0002.0001.5002.5-02.0-04.5-04.0-04.0-03.0-06.0003.5
-04.0000.0-02.5000.0-02.0-02.0-03.0-02.5-01.5-03.0002.0-03.5-04.5-05.0-05.0
003.0-03.0-04.0-07.0-01.5-02.5002.0-05.5-03.5013.0
-23.5-25.0019.5-03.0-25.0-23.5-00.5001.0-25.0-24.5005.0-01.0-25.5-25.0017.0
017.0-23.5-24.5-01.5000.0-24.0-24.0015.5015.5-23.5-23.5-01.5-00.5-24.5-25.0
019.5-02.5-25.0-25.0013.5000.0-24.5-24.5018.0000.0-24.5-24.5021.0-04.5-24.5
-25.0-10.5-02.5-24.5-25.0002.5-07.5-25.0-25.5006.5-01.0-23.5-25.5-07.5010.0
-24.0-25.0-03.0-01.0-24.5-25.0010.0-04.0-25.0-25.0-01.0002.0-24.5-25.5006.0
007.0-23.0-25.0000.0-01.5-24.0-23.5003.5001.0-25.0-25.0-02.0-06.0-18.5-28.5
001.5-03.5-24.0-24.0-02.0-04.0-24.0-24.0001.0-00.5-24.5-24.5002.0001.0-22.0
-23.0-01.5-04.0-23.5-25.0000.5-03.0-24.0-25.0000.0001.0-25.0-25.0001.0-04.5
-23.0-23.0-01.0-00.5-24.5-25.0008.5-04.5-23.0-23.5-01.5003.5-22.0-25.0002.5
001.0-24.0-25.0002.5001.5-22.0-22.0000.0002.0-24.0-23.5-05.0000.5-24.0-25.0
-02.0003.0-24.0-25.0018.0-01.0-24.5-24.5007.5000.0
003.0005.0004.5003.0003.0002.0004.5002.5002.0002.0001.0000.0-02.0000.5003.0
003.5002.5002.0001.5001.0005.5002.0002.0002.5003.0007.0005.0002.0004.5003.5
007.0003.5004.0003.5007.5006.0003.5004.5000.0001.0005.5002.5004.5004.0006.0
005.0002.5001.0001.5003.0001.0001.0001.0006.5005.5003.5003.5003.0002.5002.5
```

APPENDIX F

STATISTICAL ANALYSES

Analyses of variance (ANOVA) were administered for data from each of the four subtests of the Willeford CAP test battery (CST, LPFS, BF, and RASP) using a computer program developed by Ullrich and Pitz which was based on designs elaborated by Winer (1971). The program was processed by a DECsystem-20 computer.

Table F1 contains the ANOVA summary for the analysis related to the CST. This analysis revealed a significant interaction of Test (N = 2) by Item (N = 10) by Channel (N = 2). This interaction is illustrated in Figure 1 in the main body of the text. Although this interaction is significant beyond the 0.005 level of significance (see Table F1), the clinical usefulness of this interaction remains in doubt. Specifically, it appears that test items recorded on Channel 1 cover a slightly larger range of values than those recorded on Channel 2. The test items for Test 2, Channel 2 appear to exhibit a lesser range of values than any other condition. The mean recording levels are not significantly different clinically (less than 2.5 dB), and most of the test items are recorded at levels that are within 2.0 dB of one another, and the extreme range of discrepancy is 5.93 dB. Statiscally significant interactions of Test by Channel, Test by Item, and of Item by Channel also were demonstrated. However, the extreme ranges of the mean differences for each of these two-way interactions were 0.8 dB, 2.8 dB, and 3.3 dB respectively. These differences are not considered clinically relevant, since current audiometer standards (ANSI, 1969) allow a 2.5 dB error in calibration. The change in performance that would result from these discrepancies remains uncertain. The only main effect that was found to be significant was the Item effect which would indicate that some items are recorded at significantly different levels than other items. Examination of the data failed to reveal a consistent pattern of these recording-level discrepancies, and it has been assumed that during testing, this would become a random effect.

TABLE F1

CST: ANOVA summary table.
Group (G) ["OLD" versus "NEW"], Test (T) [Test 1 versus Test 2], Item (I) [Item 1 versus Item 2, ... , versus Item 10], Channel (C) [Channel 1 versus Channel 2].

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
G	1386.1100	1386.1100	1	1.26	0.3052
Error	6602.67	1100.45	6		
T	0.3781	0.3781	1	0.02	0.8767
Error	95.3093	15.8849	6		
GxT	1.7999	1.7999	1	0.11	0.7449
Error	95.3093	15.8849	6		
I	79.4094	8.8233	9	3.08	0.0050
Error	154.922	2.8690	54		
GxI	21.4813	2.3868	9	0.83	0.5908
Error	154.922	2.8690	54		
TxI	145.778	16.1976	9	5.76	0.0001
Error	151.784	2.8108	54		
GxTxI	14.8250	1.6472	9	0.59	0.8035
Error	151.784	2.8108	54		
C	0.0125	0.0125	1	0.00	0.9716
Error	58.0218	9.6703	6		
GxC	31.8781	31.8781	1	3.30	0.1176
Error	58.0218	9.6703	6		
TxC	39.2000	39.2000	1	26.77	0.0026
Error	8.7846	1.4641	6		

TABLE F1: (continued)

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
GxTxC	0.7031	0.7031	1	0.48	0.5189
Error	8.7846	1.4641	6		
IxC	157.2380	17.4708	9	8.88	0.0000
Error	106.1970	1.9666	54		
GxIxC	4.0281	0.4476	9	0.23	0.9884
Error	106.1970	1.9666	54		
TxIxC	96.7688	10.7521	9	3.24	0.0035
Error	179.3110	3.3206	54		
GxTxIxC	18.6094	2.0677	9	0.62	0.7735
Error	179.3110	3.3206	54		

Table F2 contains the ANOVA summary for the analyses related to the LPFS. These analyses revealed a significant List (N = 2) by Item (N = 50) interaction. This result would indicate that the recording levels of test items varied as a function of the list number (List 1 versus List 2). The range of recording levels for test items of List 1 was -9.4 dB to -2.5 dB (a range of 6.9 dB). The range of values for the test items of List 2 was -9.9 dB to 0.1 dB (a range of 10.0 dB). Although the differences in the recording levels of test items for each test appear significant beyond the 0.00001 level of significance, the amount of overlap between the levels of test items in List 1 and those of List 2 would make this statistical significance of little clinical use. These analyses also reveal that the main effect for Item indicate that some items are recorded at significantly different levels than other items. Although the range of means encompases 5 to 10 dB, the clinical relevance of this effect is questionable, since lesser ranges would be indicative of compressed speech.

TABLE F2

LPFS: ANOVA summary table.
Group (G) ["OLD" versus "NEW"], List (T) [List 1 versus List 2], Item (I) [Item 1 versus Item 2, ... , versus Item 50].

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
G	2330.7400	2330.7400	1	1.05	0.3460
Error	13279.80000	2213.3000	6		
T	1.2403	1.2403	1	0.02	0.8762
Error	310.1140	51.6857	6		
GxT	108.4130	108.4130	1	2.10	0.1963
Error	310.1140	51.6857	6		
I	1148.5800	23.4404	49	7.98	0.0000
Error	863.4210	2.9368	294		
GxI	66.7777	1.3628	49	0.46	0.9990
Error	863.4210	2.9368	294		
TxI	1313.6500	26.8092	49	8.22	0.0000
Error	959.1040	3.2623	294		
GxTxI	102.6030	2.0939	49	0.64	0.9695
Error	959.1040	3.2623	294		

Table F3 contains the ANOVA summary for the analyses related to the BF. These analyses reveal a List (N = 2) by Item (N = 20) by Channel (N = 2) by Syllable (N = 2) interaction. This interaction would then indicate that every factor considered interacts in some way with every other factor. It may be explained by virtue of the fact that the low-pass filtered items on one channel were consistently recorded at a higher level than the high-pass filtered items, and that the discrepancy between the high-pass filtered items and the low-pass filtered items was different for each tape, as well as between List 1 and List 2 of a given tape. Furthermore, the amount of discrepancy also varied with each item. For clinical purposes the analyses that appear most relevant include the lack of a significant group effect indicating that the average recording level was the same for "NEW" tapes as it was for "OLD" tapes. In addition, there was no significant list effect, which may lead one to conclude that the average recording level was the same for List 1 as it was for List 2. By contrast, there were significant effects for Item, Channel, Syllable, List by Item, List by Channel, List by Item by Channel, Item by Syllable, List by Item by Syllable, Channel by Syllable, Item by Channel by Syllable. The significant Channel effect simply indicates the obvious: that -4.1 dB is significantly different from -29.1 dB, which is to say that the high-pass channel recording level is significantly less than that of the low-pass channel. Although a significant simple effect for syllable is demonstrated beyond the 0.005 level of significance, there is only a 1.8 dB difference between the first syllable recording level and that of the second syllable. This difference was not taken as clinically relevant.

TABLE F3

BF: ANOVA summary table.
Group (G) ["OLD" versus "NEW"], List (T) [List 1 versus List 2], Item (I) [Item 1 versus Item 2, ... , versus Item 20], Channel (C) [Channel 1 versus Channel 2], Syllable (S) [Syllable 1 versus Syllable 2].

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
G	2.6281	2.6281	1	0.001	0.9740
Error	15056.2000	2509.3700	6		
T	2.3633	2.3633	1	0.09	0.7683
Error	155.2140	25.8690	6		
GxT	13.4070	13.4070	1	0.52	0.5034
Error	155.2140	25.8690	6		
I	879.4060	46.2845	19	5.89	0.0000
Error	895.4750	7.8550	114		
GxI	77.3719	4.0722	19	0.52	0.9496
Error	895.4750	7.8550	114		
TxI	591.5740	31.1355	19	2.87	0.0004
Error	1235.2500	10.8356	114		
GxTxI	35.4992	1.8684	19	0.17	0.9999
Error	1235.2500	10.8356	114		
C	199900.0000	199900.0000	1	376.39	0.0001
Error	3186.6200	531.1030	6		
GxC	1428.0500	1428.0500	1	2.69	0.1505
Error	3186.6200	531.1030	6		
TxC	102.9440	102.9440	1	1.66	0.2444
Error	372.2320	62.0387	6		

TABLE F3: (continued)

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
GxTxC	9.9741	9.9741	1	0.16	0.7020
Error	372.2320	62.0387	6		
IxC	1086.7400	57.1968	19	7.23	0.0000
Error	901.5110	7.9080	114		
GxIxC	71.6367	3.7704	19	0.48	0.9672
Error	901.5110	7.9080	114		
TxIxC	685.1800	36.0621	19	2.93	0.0004
Error	1403.3100	12.3097	114		
GxTxIxC	62.5379	3.2941	19	0.27	0.9988
Error	1403.3100	12.3097	114		
S	969.5280	969.5280	1	41.72	0.0011
Error	139.4440	23.2406	6		
GxS	0.8000	0.8000	1	0.03	0.8525
Error	139.4440	23.2406	6		
TxS	96.2508	96.2508	1	9.19	0.0227
Error	62.8517	10.4753	6		
GxTxS	4.3945	4.3945	1	0.42	0.5457
Error	62.8517	10.4753	6		
IxS	867.8630	45.6770	19	11.53	0.0000
Error	451.7120	3.9624	114		
GxIxS	71.3406	3.7548	19	0.95	0.5272
Error	451.7120	3.9624	114		

TABLE F3: (continued)

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
/	/	/	/	/	/
TxIxS	1586.8900	83.5205	19	13.48	0.0000
Error	706.117	6.1940	114		
GxTxIxS	63.0586	3.3189	19	0.54	0.9407
Error	706.117	6.1940	114		
CxS	161.0280	161.0280	1	22.99	0.0035
Error	42.0345	7.0058	6		
GxCxS	4.2743	4.2743	1	0.61	0.5309
Error	42.0345	7.0058	6		
TxCxS	92.9868	92.9868	1	3.47	0.1103
Error	161.0250	26.8375	6		
GxTxCxS	0.7545	0.7545	1	0.03	0.8662
Error	161.0250	26.8375	6		
IxCxS	620.0490	32.6341	19	6.48	0.0000
Error	574.4430	5.0390	114		
GxIxCxS	61.7734	3.2512	19	0.65	0.8636
Error	574.4430	5.0390	114		
TxIxCxS	1533.8700	80.7302	19	13.54	0.0000
Error	679.9600	5.9646	114		
GxTxIxCxS	33.2246	1.7487	19	0.29	0.9980
Error	679.9600	5.9646	114		

Table F4 contains the ANOVA summary for the analyses related to the RASP. Results from these analyses reveal that there is a significant interaction of List (N = 3) by Item (N = 10) by Channel (N = 2). This interaction would appear to indicate that recording levels for test items differ as a function of Channel the recording levels of which differ as a function of list. Examination of the data reveals a total range of mean recording level values of -0.6 dB to 6.1 dB. Therefore, the largest difference in this interaction is in the order of 6.7 dB, which may not be considered clinically relevant. In spite of the rather small total range of values, the three-way interaction was significant beyond the 0.001 level of significance. Two-way interactions of List by Item and of Item by Channel also are demonstrated beyond the 0.01 level of significance. In neither case is the mean difference amongst these variables greater than 5 dB, and there is much overlap in the values. Therefore, although statistical analyses reveal significant interactions which should alert the clinician to be aware of inter-relations amongst these variables during the administration of the lists, the clinical relevance of these interactions may be of questionable value. In addition, significant main effects of List and Item are demonstrated. The difference between mean recording levels for the three lists (across tapes) was less than 1 dB at the extreme. The difference between mean recording levels for the various test items (collapsed across all other variables) was less than 2.5 dB. However, differences between individual items within a test and/or within a channel were larger as discussed in the main text of this report.

TABLE F4

RASP: ANOVA summary table.
Group (G) ["OLD" versus "NEW"], List (T) [List 1 versus List 2 versus List 3], Item (I) [Item 1 versus Item 2, ... , versus Item 10], Channel (C) [Channel 1 versus Channel 2].

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
G	561.1690	561.1690	1	0.42	0.5444
Error	7971.4800	1328.5800	6		
T	94.3760	47.1880	2	14.68	0.0009
Error	38.5874	3.2156	12		
GxT	4.0031	2.0016	2	0.62	0.5571
Error	38.5874	3.2156	12		
I	239.0400	26.5600	9	8.85	0.0000
Error	162.1230	3.0023	54		
GxI	17.1125	1.9014	9	0.63	0.7646
Error	162.1230	3.0023	54		
TxI	169.8640	9.4369	18	2.49	0.0023
Error	409.1210	3.7882	108		
GxTxI	45.9656	2.5537	18	0.67	0.8299
Error	409.1210	3.7882	108		
C	0.1688	0.1688	1	0.01	0.9329
Error	140.6230	23.4371	6		
GxC	9.6333	9.6333	1	0.41	0.5497
Error	140.6230	23.4371	6		
TxC	11.5406	5.7703	2	0.68	0.5306
Error	102.3960	8.5330	12		

TABLE F4: (continued)

SOURCE	SUM OF SQUARES	MEAN SQUARE	DF	F-RATIO	PROB.
GxTxC	9.7635	4.8818	2	0.57	0.5833
Error	102.3960	8.5330	12		
IxC	160.8620	17.8736	9	3.02	0.0056
Error	319.3980	5.9148	54		
GxIxC	66.9812	7.4424	9	1.26	0.2803
Error	319.3980	5.9148	54		
TxIxC	176.3030	9.7946	18	2.85	0.0006
Error	371.7700	3.4423	108		
GxTxIxC	71.8095	3.9894	18	1.16	0.3079
Error	371.7700	3.4423	108		