
Evaluation of the Willeford Binaural Resynthesis Subtest

Évaluation du sous-test de resynthèse binaurale de Willeford

J.L. Stouffer

Department of Communicative Disorders

The University of Western Ontario

Carol Wilson

Mississauga Hospital

Mississauga, Ontario

Abstract

The binaural resynthesis subtest of the Willeford Test Battery (1977) was administered to a sample of 56 children with an average age of 8 years, 8 months, in order to investigate test characteristics. Analysis revealed that 19 of 40 items fell outside the optimal range for item difficulty. In addition, a significant difference was found between right and left ear scores, and a significant correlation was found between item difficulty and word frequency. Based on the results of these analyses, it was concluded that the use of the Willeford binaural resynthesis subtest as a measure of central auditory processing with children of elementary school age is questionable.

Résumé

Le sous-test de resynthèse binaurale, qui fait partie de la batterie de test de Willeford (1977), a été administré à un groupe de 56 enfants dont l'âge moyen était de 8 ans et 8 mois afin d'examiner les caractéristiques du test. L'analyse effectuée a démontré que 19 des 40 mots étudiés ne correspondaient pas au niveau de difficulté théoriquement prévu. De plus, on a constaté une différence significative entre les résultats de l'oreille gauche et ceux de l'oreille droite et un rapport certain entre le niveau de difficulté et la fréquence des mots. Les résultats de ces analyses permettent de conclure que l'utilisation du sous-test de resynthèse binaurale de Willeford comme mesure de l'audition centrale des enfants d'âge élémentaire est discutable.

Evaluation of the Willeford Binaural Resynthesis Subtest

Central auditory tests were designed for the purpose of identifying brainstem and cortical pathologies. The tests were based on the idea that reducing the redundancy of a complex signal, such as speech, would hinder test performance in persons with central auditory lesions, regardless of peripheral hearing sensitivity (Martin & Clark, 1977; Willeford & Billger, 1978). More recently, interest has focused on the application of central auditory tests to the school-age population (Ivey, 1986) in an attempt to explain academic under-achievement and behaviour problems. Although central auditory tests are widely used

with children, few normative data exist concerning test characteristics.

Matzker (1959) introduced a test that included the principles of binaural resynthesis. His procedure consisted of filtering spondaic words into low and high frequency segments, to be simultaneously presented to opposite ears. Neither segment alone provided sufficient information to be intelligible for normal hearing adult subjects; however, intelligibility improved considerably when the segments were presented together in a dichotic fashion. Matzker suggested that the resynthesis of these segments occurred in the brainstem. Matzker's theory that binaural resynthesis is a sensitive test of brainstem integrity has received some empirical support (Bocca & Calero, 1963; Dempsey, 1978, 1983; Matzker, 1959). Other authors, however, have suggested that the filtering of test items reduces the redundancy of the speech signal to the extent that some cortical integration is also required of the listener (Dempsey, 1977; Linden, 1964; Protti, 1983; Roush & Tait, 1984).

Based on Matzker's model, several binaural resynthesis tests have been developed for clinical or experimental use with adults and/or children (Linden, 1964; Martin & Clark, 1977; Musiek, Geurkink, & Kietel, 1982; Palva & Jokinen, 1975; Plakke, Orchik, & Beasley, 1981; Roush & Tait, 1984; Smith & Resnick, 1972; Willeford, 1977). These binaural resynthesis tests involve either monosyllabic words or spondaes, varying high and low frequency bandwidths, and/or varying presentation levels. The test of interest in the present study is the binaural resynthesis subtest in the Willeford Test Battery (Willeford, 1977). Willeford's binaural resynthesis subtest was originally constructed as a site-of-lesion test for use with an adult population. Later, White (1977) collected data on children. A review of White's (1977) data reveals a slight maturation effect and a tendency for binaural resynthesis abilities to be highly variable in normal young children. Willeford's binaural resynthesis subtest consists of two lists of 20 spondaic words that have been filtered into high frequency bandpass (1900-2100 Hz) and low frequency bandpass (500-700 Hz) segments. For the first list presented, the low-

pass segment is presented to one ear (e.g., the right), while the opposite ear (e.g., the left) simultaneously receives the high-pass segment. The conditions are reversed for the second list, and scores for each list are referenced to the ear receiving the low frequency portion (Dempsey, 1978). There is no standard presentation level for the binaural resynthesis subtest; however, test results are dependent on the presentation level (Musiek & Geurkink, 1980; Musiek et al., 1982; White, 1977).

Willeford's binaural resynthesis subtest continues to be administered in two parts (i.e., relative to the right and left ear) despite the fact that research has not provided any rationale for determining separate ear measures. It has been reasoned that binaural resynthesis, if it occurs, should do so regardless of the ear receiving the high and low frequency portions (Dempsey, 1978). However, when the Willeford binaural resynthesis subtest is administered to children, ear differences are sometimes demonstrated (e.g., White, 1977). One possible reason for such differences is that the two lists are not equally difficult (Dempsey, 1978).

A possible imbalance in the difficulty of the two Willeford binaural resynthesis lists may be related to the level of vocabulary used in the test. Willeford's (1977) criteria for selecting his binaural resynthesis test words were that neither the high bandpass nor the low bandpass component alone provided sufficient acoustic information to allow the word to be identified by normal adult listeners. Although the vocabulary was judged to be appropriate for adults, Dempsey (1983) questions its application to a much younger population and points out the need for vocabulary to be appropriate to the child's age and experience level.

In addition to the questionable vocabulary noted by Dempsey (1983), it has been demonstrated that the frequency of occurrence of words (word frequency) affects auditory recognition, with frequently occurring words more easily recognized than infrequently occurring words (Paivio & Begg, 1981). Word frequency was not considered by Willeford (1977).

McCauley and Swisher (1984) note that the validity of a test is jeopardized whenever extraneous variables are involved and the test fails to measure only the behaviour of interest. Clearly, word frequency or difficult vocabulary could confound the binaural resynthesis task, particularly for a language-learning disabled child who is being assessed for a suspected auditory perceptual deficit.

White (1977) administered Willeford's binaural resynthesis subtest to a group of normal children (including 8 and 9 year old children) using two different presentation levels, 30 dB and 40 dB SL. She found increased performance and a decrease in variability when the higher presentation level was

used. These findings led her to use this level routinely in clinical application of the binaural resynthesis subtest. Although many clinicians may have followed White's suggestions, simply presenting the binaural resynthesis subtest at higher presentation levels will not eliminate the possible problem of unequal item difficulty (proportion correct scores). Combining test items with high and low item difficulty will result in smaller test variability than predicted by the number of test items and the item difficulty level (Hagerman, 1976), but test sensitivity is adversely affected. Dillon (1983) pointed out that test sensitivity is maximized if all items are equally difficult. However, he questioned whether equal item difficulty was practical and suggested that test items should be of similar difficulty regardless of the optimum item difficulty. Thus, test item difficulty must be considered in relation to test sensitivity as well as test reliability.

Item sensitivity is frequently misunderstood or ignored. An item with maximum sensitivity will have an item difficulty level of 0.5 (Anastasi, 1982). A test constructed with all items having an item difficulty level of 0.5 will result in an overall test score of 50%. If a test were constructed so that half the words were always correctly identified and the other half were always incorrectly identified, a score of 50% would also be obtained. Although the overall score is 50%, the test has zero sensitivity since all of the test items are either too easy or too difficult (ceiling and floor effects). Dillon (1983) provides a hypothetical example of how excessive variation in item difficulty can obscure results. In his example, two test scores of 33/50 and 39/50 are obtained which are not significantly different at the 95% confidence level. However, upon closer examination, we find that in each list 30 words are always correct and 10 words are always incorrect with 3/10 and 9/10 remaining words resulting in the differences between the two lists. This significant difference (30% versus 90%) was obscured by the other 40 words, which were too easy or too difficult. The penalty for using such an inappropriate test is paid in two ways (Dillon, 1983). First, the time spent administering the noninformative items is wasted. Second, the presence of noninformative items obscures what would have been a significant difference in performance.

As the proportion of correct responses for an item approaches 0.00 or 1.00, the sensitivity of the item approaches zero; of course, as sensitivity decreases, reliability increases and vice versa. Although the most sensitive proportion correct score is 0.5 for many tests (Anastasi, 1982), in the case of tests used to classify results into normal and abnormal, an item difficulty level of between 0.8 and 0.9 for the normal subjects is appropriate (Dillon, 1983). This level is away from the upper and lower limits, that is, the boundaries of too easy and too hard. Item analysis is an accepted means for determining the appropriateness of test items (Anastasi, 1982; McCauley & Swisher, 1984).

Until recently, audiology graduates from the University of Western Ontario were instructed in the use of the Willeford binaural resynthesis test for assessing a child's CAP ability. Thus, the Willeford binaural resynthesis test is in use throughout Canada, and its use is particularly widespread in Ontario. The purpose of this study was to examine the difficulty of the test items contained in Willeford's binaural resynthesis subtest and to examine the relationship between word frequency and item difficulty. This information would indicate the appropriateness of the stimuli used in the Willeford binaural resynthesis subtest. In addition, the study was designed to determine whether differences in children's performance could be demonstrated as a function of the ear being referenced. This information could indicate whether or not some degree of cortical integration was involved in the binaural resynthesis task.

Method

Subjects

The sample consisted of 26 male and 30 female children from age 8 years, 1 month to 9 years, 3 months ($M = 8$ years, 8 months). Subjects were selected from grade 2 and 3 classes in local schools on the basis of teacher reports that each child had average or above average academic achievement in all areas. It was assumed that children with no record of previous academic failure who were currently obtaining passing grades

would not have significant learning disabilities or neurological deficits. All potential subjects had their hearing screened at 20 dB HL (Re: ANSI S-3.6 1969) for 1000, 2000, and 4000 Hz. Immittance screening was also administered to rule out the presence of middle ear dysfunction. Pass criteria for impedance screening was a tympanometric configuration with a definite pressure peak between -100 and +100 mm H₂O. To ensure that limited vocabulary did not affect test results, subjects were given the *Peabody Picture Vocabulary Test-Revised* [PPVT-R] (Dunn & Dunn, 1981) and the *Expressive One-Word Picture Vocabulary Test* [EOWPVT] (Gardner, 1979) as measures of receptive and expressive vocabulary, respectively. Successful candidates for the study obtained PPVT-R and EOWPVT scores at or above the 22 percentile. All initial testing was completed in school settings.

Procedure

On the day of experimental testing, subjects were required to demonstrate pure tone air conduction thresholds of 15 dB HL or better (re ANSI S3.6 1969) at 500, 1000, and 2000 Hz, and speech discrimination scores of 90% or better for Phonetically Balanced Kindergarten (PBK) words presented at 30 dB above speech reception threshold. All audiological testing, including Willeford's binaural resynthesis subtest, was administered in a double-walled sound booth meeting ANSI specifications.

The recorded spondees¹ in the binaural resynthesis sub-test were presented through headphones via a two-channel Luxman K-117 cassette player connected to a two-channel clinical audiometer (Grason Stadler GSI 10). In order to eliminate the possible effects of unequal list difficulty on results, the two 20-word binaural resynthesis lists were combined to develop two word orders of forty words each (List 1 first combined with List 2 second= I-40, or List 2 first combined with List 1 second= II-40). Prior to administering the binaural resynthesis subtest, the words were read in random order to the subject to familiarize him/her with all 40 words.² Subjects were instructed to repeat each test item. The entire test, consisting of forty words, was then presented to each subject. The low-pass segment was delivered to either the right or left ear at 30 dB above the 500 Hz pure tone threshold for that ear; the opposite ear received the high-pass segment at 30 dB above the 2000 Hz hearing level for that ear. The only exception to this test procedure was that presentation levels were never less than 30 dB HL, regardless of hearing threshold levels. Either List I-40 or II-40 was presented first, and low-pass and high-pass segments were presented via channel A and channel B of the audiometer, respectively. By convention, the ear receiving the low portion was termed the test ear. Subjects were randomly assigned to the four resulting test conditions (Right I-40; Right II-40; Left I-40; Left II-40).

1 A copy of the Willeford Binaural Resynthesis tape was used to construct the present test tape. The tape was played back on a TEAC A-7300 two-track reel-to-reel tape recorder. Peak intensity levels were measured for each channel by delivering the recorder output to a B & K graphic level recorder (model 2307). Relative intensity levels for each word were calculated from the graphic level recordings, and the word with the lowest peak intensity was assigned a value of 0 dB. All other words were assigned a dB value greater than the reference word, and these values were used as attenuation values to equalize the peak intensity of each word in a rerecording of the word lists. During rerecording, one channel of the tape output from the attenuator was filtered to produce a 500 to 700 Hz low frequency bandpass and the opposite channel was filtered to produce a 1900 to 2100 Hz high frequency bandpass. Approximately five seconds of silence was left between each word on the tape to allow for a subject's response. The output of the filters was recorded onto a Sony LNX 30 cassette tape using a Sony TCFX 35 cassette deck. A calibration tone was recorded onto the tape using a Wavetek model 135 sine generator.

2 Familiarizing the subjects with the words may have reduced the effect of vocabulary ability and word frequency on test performance as subjects could now guess from a limited set of possible words. However, since this type of familiarization is practised in many clinical settings, it was considered to be appropriate.

In addition, 32 subjects, equally distributed over the 4 presentation groups, were randomly selected to receive a second presentation of the binaural resynthesis subtest to determine if ear differences were present. During the second test presentation, the ears receiving the low-pass and high-pass conditions were reversed.

Results

Item Difficulty

For each subject, responses to the 40 test items were scored as either correct or incorrect. The proportion of correct responses across all 56 subjects was then computed, based on the first or only test presentation. An item analysis was conducted to determine if the proportion correct score for each word (item difficulty) was significantly different from 0.85. The 0.85 level was based on the recommendations of Dillon (1983).

Table 1 lists the frequency of occurrence and the proportion correct score for each word. The proportion of correct responses is compared to 0.85, and items that are significantly different from 0.85 are listed in columns 4 and 8. In total, 19 of the 40 proportions differed significantly from 0.85; 17 were significant at or beyond the 0.01 level and 2 were significant at the 0.05 level of confidence. Thus, 19 of the test items are either too easy or too difficult.

Word Frequency

The results were analyzed to determine if a correlation existed between the proportion of time an item was correctly identified (item difficulty score) and word frequency (Carroll, Davies, & Richman, 1971).³ Item difficulty and word frequency were positively correlated ($r = 0.32, p = 0.02$) indicating that frequently occurring words were identified more often than infrequently occurring words.⁴

3 Carroll, Davies, & Richman (1971) examined the frequency of occurrence of over 5 million words to which students in grades 3 through 9 are exposed. Words were compiled from over 1,000 publications. This work was completed to develop a citation base for *The American Heritage School Dictionary*.

4 Fifteen of the forty binaural resynthesis words had word frequencies of 0. Therefore, a numerical value of 1 was added to each word frequency and a square root transformation was performed on the adjusted values prior to computing the correlation.

Ear differences

For the 32 subjects tested twice, a percent correct score for 40 items was computed for the right and left ear presentations. The mean difference between left ear ($M = 70.6\%$) and right ear ($M = 75.6\%$) scores was 5.0, with the left ear score poorer than the right ear score, $t(31) = -2.51, p < 0.02$.

Discussion

Item Difficulty

A 0.85 proportion correct score was selected to represent optimal item difficulty (Dillon, 1983). Based on this criterion level, the sensitivity of the binaural resynthesis subtest, as it is presently structured, appears to be less than optimal for use with children. Nineteen of the 40 words contained in Willeford's binaural resynthesis subtest had a significantly high or low item difficulty for 8-9 year old children (Table 1). The easiest item was eyebrow (100% correct), and the most difficult was dovetail (16% correct). Since the present sample consisted of normal children (based on expressive and receptive vocabulary measures, hearing sensitivity, and teachers' reports) a high performance rate (away from ceiling effects) was expected. As can be seen from an examination of Table 1, however, 14 words were identified at proportions significantly lower than 0.85; 5 were identified at proportions significantly higher than 0.85. Thus, for the Willeford binaural resynthesis subtest we are wasting our time by administering 19 noninformative items, and further, the use of these 19 noninformative items might obscure significant differences in performance.

One explanation for the wide range of item difficulty is that during test construction words were significantly degraded by the filtering process. One must bear in mind that binaural resynthesis may not only involve fusion of two complementary fragments of information, but also some degree of auditory closure. Therefore, a child, upon hearing the degraded signal for a word outside his/her vocabulary, might respond by substituting a familiar, similar sounding word.

The correlation between word frequency and item difficulty was significant ($p = 0.02$), but the correlation was low ($r = 0.32$). However, because all children in the present study had normal or above normal vocabulary development, as measured by the PPVT-R and the EOWPVT, a higher correlation would possibly be obtained for a population with a wider range of vocabulary skills. Also, spoken word frequency measures would probably be a more appropriate measure of the relationship between binaural resynthesis and word frequency than printed word frequency.

Table 1. Binaural resynthesis words, their frequency of occurrence and the proportion each word was correctly identified.

Item 1	Word Freq. 2	Prop. 3	Sig. Level 4	Item 5	Word Freq. 6	Prop. 7	Sig. Level 8
eyebrow	1	1.00	**	daylight	19	.77	—
rainbow	14	.98	**	although	63	.75	—
scarecrow	20	.98	**	bobwhite	0	.75	—
baseball	68	.95	**	meatball	10	.75	—
wildcat	3	.95	**	watchword	0	.71	—
churchbell	0	.93	—	dollhouse	0	.71	—
bluejay	2	.93	—	buckwheat	7	.70	*
horseshoe	18	.93	—	stairway	7	.70	*
footstool	0	.91	—	doormat	0	.66	**
shoelace	0	.91	—	therefore	13	.63	**
platform	17	.89	—	padlock	0	.59	**
workshop	2	.88	—	soybean	1	.54	**
bagpipe	1	.86	—	lifeboat	0	.52	**
bloodhound	0	.84	—	wigwam	4	.52	**
birdnest	0	.84	—	woodchuck	3	.48	**
northwest	13	.84	—	whizbang	0	.45	**
bonbon	1	.82	—	nutmeg	0	.43	**
bedroom	31	.82	—	mishap	0	.29	**
drugstore	17	.80	—	housework	3	.25	**
yardstick	15	.77	—	dovetail	0	.16	**

* Significant at the 0.05 Level
 ** Significant at the 0.01 Level Prop. = Proportion Correct

Note. Word frequency (Carroll, Davies, & Richman, 1971) for each binaural resynthesis test item is listed in columns 2 and 6. Columns 3 and 7 indicate the proportion of correct responses; starred items in columns 4 and 8 indicate proportions that are significantly different from a proportion of 0.85.

Ear Differences

Left ear scores (70.6%) were found to be significantly lower than right ear scores (75.6%). Although a difference of approximately 2 items in 40 would not be judged to be clinically significant by most audiologists, all 40 items are not contributing equally to the overall binaural resynthesis score. Based on an 0.85 item difficulty level, only 21 items represent an appropriate difficulty level. That is, the difference between ears is significant even when approximately half of the test items tend to obscure test results (see Item Difficulty above).

Willeford and Burleigh (1985) reported differences of less than 3% between right and left ear scores for children 6 to 10 years of age. White (1977) reported a range of ear asymmetry from 0% to 30%. She concluded that these differences were largely due to the imbalance of difficulty between List 1 and List 2 because List 2 was found to yield a poorer score whether it was given to the right or left ear. Right and left ear scores differed very little when compared for the same list. If binaural resynthesis is primarily a task of fusion of signals in the brainstem (Bocca & Calearo, 1963; Dempsey,

1978; Dempsey, 1983; Matzker, 1959), fusion should not depend on which ear receives the low-pass or high-pass segment. Therefore, if asymmetry between ears occurs, the only possible explanations appear to be based on differences in list difficulty or on the hypothesis that binaural resynthesis requires some degree of cortical integration (Dempsey, 1977; Linden, 1964; Protti, 1983; Roush & Tait, 1984). In the present study, List 1 and List 2 were combined to control the list difficulty and all 40 words were presented to each ear; however, a significant ear difference was still found. Therefore, the present results support the hypothesis that binaural resynthesis requires some degree of cortical integration.

Summary

In summary, left ear scores were found to be significantly lower than right ear scores; word frequency was positively correlated with item difficulty; and the item difficulty for 19 of the 40 test words was either too high or too low. Therefore, use of the binaural resynthesis subtest with children, as it is presently structured, is questionable.

Acknowledgement

This research was supported in part by a grant from Unitron Corporation. We wish to thank Mrs. Doreen Mahoney of the Catholic Schools of London and Middlesex for her cooperation in providing subjects for this research. Thanks are also extended to Dr. Carla J. Johnson for her helpful comments on earlier versions of this article.

Address all correspondence to: Dr. James L. Stouffer, Department of Communicative Disorders, Elborn College, University of Western Ontario, London, Ontario, Canada, N6G 1H1.

References

- American National Standard Specifications for Audiometers (1970). (ANSI S3.6-1969) New York: American National Standards Institute, Inc.
- Anastasi, A. (1982). *Psychological testing*. New York: MacMillan Publishing Co., Inc.
- Bocca, E., & Calearo, C. (1963). Central hearing processes. In J. Jerger (Ed.), *Modern developments in audiology*. New York: Academic Press.
- Carroll, J.B., Davies, P., & Richman, B. (1971). *Word frequency book*. New York: American Heritage Publishing Co.
- Dempsey, C. (1977). Some thoughts concerning alternate explanations of central auditory test results. In R. Keith (Ed.), *Central auditory dysfunction*. New York: Grune and Stratton.
- Dempsey, C. (1978). A guide to tests of central auditory function. *Otolaryngology Clinics of North America*, 11, 677-696.
- Dempsey, C. (1983). Selecting tests of auditory function in children. In E.Z. Lasky & J. Katz (Eds.), *Central auditory processing disorders: Problems of speech, language and learning* (pp. 203-221). Baltimore, MD: University Park Press.
- Dillon, H. (1983). The effect of test difficulty on the sensitivity of speech discrimination tests. *Journal of the Acoustical Society of America*, 73, 336-344.
- Dunn, L.M., & Dunn, L.M. (1981). *Peabody Picture Vocabulary Test Revised*. Circle Pines, MN: American Guidance Service.
- Gardner, M.F. (1979). *Expressive One-Word Picture Vocabulary Test*. Novato, CA: Academic Therapy Productions.
- Hagerman, B. (1976). Reliability in the determination of speech discrimination. *Scandinavian Audiological*, 5, 219-228.
- Ivey, R.G. (1986). Survey of school age children referred for central auditory assessment. *Human Communication Canada*, 10, 5-10.
- Linden, A. (1964). Distorted speech and binaural resynthesis tests. *Acta Oto-Laryngologica*, 58, 32-48.
- Martin, F.N., & Clark, J.G. (1977). Audiologic detection of auditory processing disorders in children. *Journal of the American Audiology Society*, 3, 140-146.
- Matzker, J. (1959). Two new methods for the assessment of central auditory functions in cases of brain disease. *Annals of Otolaryngology and Rhinology and Laryngology*, 68, 1185-1197.
- McCauley, R.J., & Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders*, 49, 34-42.
- Musiek, F.E., & Geurkink, N.A. (1980). Auditory perceptual problems in children: Considerations for the otolaryngologist and audiologist. *Laryngoscope*, 90, 962-971.
- Musiek, F.E., Geurkink, N.A., & Kietel, S.A. (1982). Test battery assessment of auditory perceptual dysfunction in children. *Laryngoscope*, 92, 251-257.
- Palva, A., & Jokinen, K. (1975). The role of the binaural test in filtered speech audiometry. *Acta Oto-Laryngologica*, 79, 310-314.
- Paivio, A., & Begg, I. (1981). *Psychology of language*. Englewood Cliffs, NJ: Prentice Hall.
- Plakke, B.L., Orchik, D.J., & Beasley, D.S. (1981). Children's performance on a binaural fusion task. *Journal of Speech and Hearing Research*, 24, 520-525.
- Protti, E. (1983). Brainstem auditory pathways and auditory processing disorders: diagnostic implications of subjective and objective tests. In E.Z. Lasky & J. Katz (Eds.), *Central auditory processing disorders: Problems of speech, language and learning* (pp. 117-139). Baltimore, MD: University Park Press.
- Roush, J., & Tait, C.A. (1984). Binaural fusion, masking level differences, and auditory brainstem responses in children with language-learning disabilities. *Ear and Hearing*, 5, 37-41.
- Smith, B.B., & Resnick, D.M. (1972). An auditory test for assessing brain stem integrity: Preliminary report. *Laryngoscope*, 82, 414-424.
- White, E.J. (1977). Children's performance on the SSW test and Willeford battery: interim clinical data. In R. Keith (Ed.), *Central auditory dysfunction* (pp. 319-340). New York: Grune and Stratton.
- Willeford, J. (1977). Assessing central auditory behaviour in children: a test battery approach. In R. Keith (Ed.), *Central auditory dysfunction* (pp. 43-68). New York: Grune and Stratton.
- Willeford, J.A., & Billger, J.M. (1978). Auditory perception in children with learning disabilities. In J. Katz (Ed.), *Handbook of clinical audiology* (pp. 410-425). Baltimore, MD: Williams and Wilkins.
- Willeford, J.A., & Burleigh, J.M. (1985). *Handbook of central auditory processing disorders in children*. Orlando, FL: Grune and Stratton.