J Am Acad Audiol 12: 7-14 (2001)

Relation Between Slopes of Word Recognition Psychometric Functions and Homogeneity of the Stimulus Materials

Richard H. Wilson*
Anne Strouse Carter*

Abstract

This tutorial paper examines the relation between the slope of a mean word recognition function and the homogeneity or variability (with respect to recognition) of the individual stimulus items that compose the test materials. This was studied in terms of both the location (Cartesian) and slope of the psychometric functions of the individual words that compose the materials. Word recognition performances were measured for 100 CID W-22 (Hirsh) words and 100 PB-50 (Rush Hughes) words in quiet (0 to 56 dB HL in 8-dB steps) on 12 subjects with normal hearing. The functions for the individual W-22 words were more homogeneous (less variable) than were the functions for the individual PB-50 words. The mean function for the W-22 words was steeper (3.1%/dB) than the function for the PB-50 words (1.9%/dB). This evaluation of the individual words demonstrates the direct relation between variability of the test items and the slope of the mean psychometric function. The more homogeneous performance is on the individual test items with respect to both location and slope, the steeper the slope of the mean psychometric function.

Key Words: Homogeneity, psychometric function, slope, variability, word recognition performance

Abbreviations: ANOVA = analysis of variance, CD = compact disc

Psychometric functions for word recognition tasks reflect the ability of a listener to understand a given set of speech materials (dependent variable) as a function of either the presentation level of the speech material or the signal-to-noise ratio of the stimulus material and a masking agent (independent variable). The ability of the listener to understand speech usually is expressed in percent correct recognition, whereas the presentation level is expressed in decibels (sound pressure level, hearing level, or signal-to-noise ratio). Two characteristics can be used to describe a word recognition psychometric function. The first is the location of the function in the Cartesian coordinate system in which the dependent variables are plotted on the y coordinate (distance from the x axis) and the independent variables are plotted on the x coordinate (distance from the y axis). The second characteristic is the slope of the function, which can be considered a special case of location. The slope of a word recognition function expresses the relation between the change in correct recognition performance (Δy) and the change in the presentation level of the signal (Δx) that is expressed as some form of $\Delta y/\Delta x$ (%/dB).

Most investigations agree that the psychometric characteristics of a set of word recognition materials are unique only for a particular recorded version of those materials (Silverman and Hirsh, 1955; Kruel et al, 1969). What, then, are the underlying mechanisms that account for the differences observed in the psychometric functions for the various versions of word recognition materials? Because recognition performance is expressed as a function of presentation level, two primary independent variables can influence the data. The first variable is the listeners from whom the data are obtained. For

*Auditory Research Laboratory, James H. Quillen VA Medical Center, Mountain Home, Tennessee and Departments of Surgery and Communicative Disorders, East Tennessee State University, Johnson City, Tennessee
Reprint requests: Richard H. Wilson, Audiology—126, VA Medical Center, Mountain Home, TN 37684

example, several studies have shown that the mean word recognition functions for listeners with hearing impairments are displaced (in location) to the higher presentation levels and have more gradual slopes when compared to the functions for the same materials obtained from young adults with normal hearing (Tillman and Carhart, 1966; Kopra et al, 1968; Wilson et al, 1976; Beattie, 1989).

The second variable is the speech stimulus, which includes the word recognition material, the particular recording (speaker) of that material, and the calibration of the material. Factors to consider with the word recognition materials are how familiar the words are to the listener, how redundant the words are (e.g., monosyllabic vs spondaic words), and how easily the target word can be confused with other words (Miller et al, 1951; Luce, 1986; Luce and Pisoni, 1998). As has been demonstrated by numerous investigations of word recognition performance, different materials spoken by different speakers and even the same materials spoken by different speakers produce psychometric functions that are different in their location and slope characteristics (Silverman and Hirsh, 1955; Tillman and Carhart, 1966; Rintelmann et al, 1974; Wilson et al, 1990; Heckendorf et al, 1997). Finally, the method of calibration is a determinant of the location of a psychometric function in the x-axis domain. Heckendorf et al (1997) showed that functions for the Hirsh recordings of the CID W-22 words were displaced about 10 dB, with one study calibrating to the average levels of the materials (Hirsh et al, 1952) and the other study calibrating to the peak levels of the materials.

Underlying the speech stimulus variable is the recognition homogeneity (or recognition variability) of the location and slope of the functions for the individual words that compose the set of materials. In an early experiment, Harris (1948) adjusted the levels of the psychoacoustic laboratory PB-50 words to equate the words in terms of recognition. The psychometric function for the adjusted or "re-recorded" words was steeper than the function for the words recorded at the original levels. In an attempt to make the thresholds of the CID W-1 spondaic words more homogeneous, Hirsh et al (1952) increased or decreased the levels of selected words before the words were recorded in their final format. In both of these experiments, the locations of the psychometric functions for the individual test items were manipulated by changing the presentation levels. The changes in presentation levels made the functions for the individual test words more homogeneous, which, in turn, produced a steeper mean function. The manipulations in these experiments basically produced a dc shift in the locations of the word functions; the slopes of the word functions, however, were not altered. Hirsh (1952) observed that the steepness of psychometric functions for words was representative of the homogeneity of the speech materials. Hirsh suggested that one reason the function for spondaic words was steeper than the function for PB-50 words was because the spondaic words were the more homogeneous material. Simulations of psychometric functions demonstrate the direct relation between the slope of the mean psychometric function for a word list and the underlying homogeneity of the psychometric functions for the individual words that compose the word list (Wilson and Margolis, 1983, Fig. 5-3). The simulations illustrated that the more homogeneous the functions of the individual words were with respect to location and slope, the steeper the slope of the mean function.

The primary purpose of this tutorial report was to demonstrate empirically the relationship between the homogeneity of the individual word functions and the mean function for that set of materials. That is, how do the locations and slopes of the functions for the individual words influence the mean function for those words? This was accomplished by establishing psychometric functions for each stimulus word that comprised two lists of the CID W-22 materials recorded by Hirsh et al (Hirsh et al, 1952) and two lists of the PB-50 materials recorded by Rush Hughes. Although the principles of stimulus homogeneity apply to any set of word recognition materials, the Hirsh W-22 and the Rush Hughes PB-50 materials were selected because recent data demonstrate that the mean psychometric functions for the two materials have substantially different slope characteristics of 3.1 percent/dB and 1.7 percent/dB, respectively (Heckendorf et al, 1997). The slopes of the mean functions suggest that the functions for the individual W-22 words should be more homogeneous (less variable) with respect to location and slope than the functions for the individual PB-50 words.

As slopes of functions were one of the foci of this report, the methods used to determine the slopes of the functions also were of interest. Although mean psychometric functions for word recognition materials are described as sigmoid shaped, the slopes of these functions, like those reported by Heckendorf et al (1997), typically are expressed as a linear slope ($\Delta y/\Delta x$) calculated between the 20 percent and 80 percent correct points. The reason for using a linear model to describe a curvilinear word recognition function is steeped in tradition that began when mathematical descriptions of data were beyond the capabilities of most laboratories (i.e., before computers). The reasoning was that the curvilinear functions "appeared" to be "linear" between the 20 percent and 80 percent correct points. Estimates of the slopes were determined visually, often with a pencil and ruler (Hirsh et al, 1952; Tillman and Carhart, 1966). As computer technology has become commonplace, we are able to describe data mathematically in an expeditious manner and to examine precisely aspects of data such as the slopes of functions at specified points. Therefore, a secondary purpose of this report was to examine the relationships among the slopes of word recognition functions calculated with various strategies.

METHODS

PB-50 lists 8 and 9 recorded by Rush Hughes, CID W-22 lists 3 and 4 recorded by Hirsh, and a 1000-Hz calibration were copied digitally (Pinnacle, Model RCD 1000) from the *Speech Recognition and Identification Materials, Disc 1.1* compact disc (CD) produced by the Department of Veterans Affairs (1991). For the PB-50 lists, the carrier phrases (*Number one* to *Number fifty*) and words were edited into 150 separate files (50 carrier phrases and 100 words). For the W-22 lists, the words with their respective carrier phrases (*You will say. . .*) were edited into 100 individual files. Seven randomizations of each of the four word lists were recorded on CD in 150-s blocks of 25 words. Interstimulus intervals of 4.5 s were used with both sets of materials. With the PB-50 lists, the carrier phrases (*Number one. . .*, etc.) were concatenated to the words as the sequence progressed through each randomization (i.e., each word had a different carrier phrase in each randomization).

Based on earlier data (Heckendorf et al, 1997), eight presentation levels in 8-dB steps between 0 and 56 dB HL were used. Because there were eight levels and only seven randomizations of each list (owing to time on the CD), the same randomizations were used at 0 and 56 dB HL. Each subject listened to a randomization of each of the four 50-word lists at each of the eight presentation levels. The 32 50-word lists were divided into 64 25-word lists. To minimize the effects of word/task familiarity and listener fatigue, the order of these 64 25-word lists was randomized for each subject.

Twelve young adults (mean age = 24.6 years) with normal hearing served in two 75-minute sessions during each of which 32 25-word lists were presented. The materials were reproduced on a CD player (Sony, Model CDP-497), amplified/attenuated (Grason Stadler, Model 10), and presented monaurally through a TDH-50P earphone encased in a P/N 510C017-1 cushion. The subjects responded verbally. The subject responses as perceived by the tester were recorded onto a spreadsheet that was used to generate data for each subject and each word. The W-22 and PB-50 data for each of the 12 subjects and the subject data for each of the 200 words then were fit with third-degree polynomials. The polynomial equations, which were used simply to describe the data with no modeling implied, were used to calculate locations and slopes of the functions at selected intervals (typically 10 percent correct intervals between 10 percent correct and 90 percent correct).

RESULTS AND DISCUSSION

The mean data for W-22 lists 3 and 4 (circles) and for PB-50 lists 8 and 9 (squares) are shown in the top panel of Figure 1 and are listed (with standard deviations by subject) in Table 1. The lines connecting the datum points are the best-fit, third-degree polynomials. A list by subjects, repeated-measures analysis of variance (ANOVA) revealed no significant main effect for PB-50 lists 8 and 9 (F [1,11] = 0.69, $p > .05$) or for W-22 lists 3 and 4 (F [1,11] = 0.01, $p > .05$). Thus, the word list data for the two lists of each material were combined for the remainder of the analysis. The combined data from the two lists are depicted in the bottom panel of Figure 1 along with similar data on the two materials from Heckendorf et al (1997). Although the current study used two lists of the two materials and Heckendorf et al used four lists of the two materials, the corresponding functions from the two studies are in good agreement in terms of both location and slope.

The influence of (1) stimulus material (W-22/PB-50) and (2) presentation level (0–56 dB HL in 8-dB steps) was assessed using a repeated-measures ANOVA. The ANOVA revealed that (1) the recognition data for the W-22 materials were significantly different from the recognition data for the PB-50 materials (F [1,11] = 89.7, $p < .0001$) and (2) the recognition performances at the various presentation

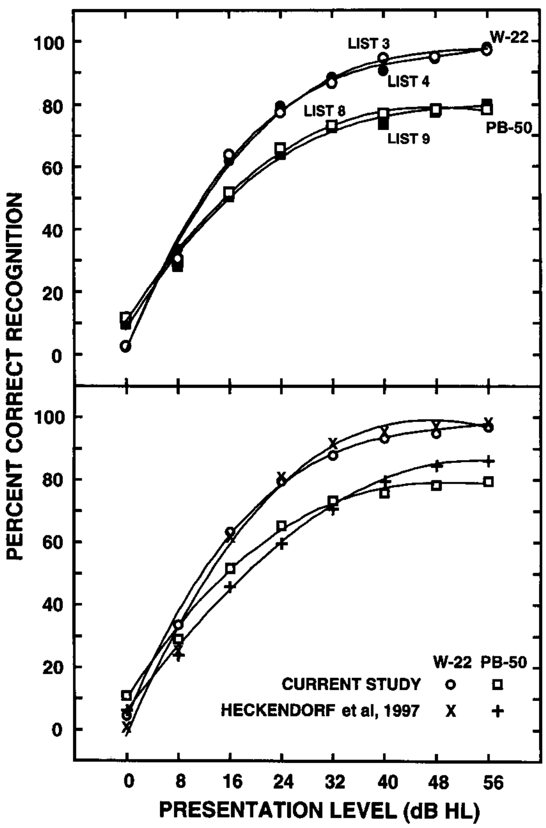

Figure 1 *Top panel:* The mean psychometric functions for the W-22 materials (lists 3 and 4, circles) and the PB-50 materials (lists 8 and 9, squares) from 12 adults with normal hearing. The lines connecting the datum points are the best-fit, third-degree polynomials (W-22, list 3: $y = 1.1333 + 4.9555x - 0.0863x^2 + 0.0005x^3$ and list 4: $1.0576 + 5.2297x - 0.1011x^2 + 0.0007x^3$; PB-50, list 8: $y = 10.3364 + 3.2199x - 0.0439x^2 + 0.0001x^3$ and list 9: $y = 8.0258 + 3.4089x - 0.0540x^2 + 0.0003x^3$). *Bottom panel:* The mean psychometric functions for the W-22 and PB-50 materials from the current study (circles and squares) and from the Heckendorf et al (1997) study using the same materials (crosses and plus symbols). The lines connecting the datum points are the best-fit, third-degree polynomials (current study, W-22: $y = 2.8803 + 4.9917x - 0.0917x^2 + 0.0006x^3$ and PB-50: $y = 9.2288 + 3.3070x - 0.0487x^2 + 0.0002x^3$; Heckendorf et al, W-22: $y = -2.2606 + 4.8894x - 0.0711x^2 + 0.0003x^3$ and PB-50: $y = 5.5000 + 2.7809x - 0.0206x^2 + 0.0001x^3$).

levels were significantly different ($F[7,77] = 241.7$, $p < .0001$). Further analysis using paired t-tests indicated that there was a significant difference in performance on the W-22 and PB-50 word lists at seven of the eight presentation levels ($p > .05$); the exception was the 8-dB HL presentation level ($p < .05$). There was a significant interaction of stimulus material and presentation level ($F[7,77] = 13.95$, $p < .0001$). As displayed in Figure 1, this interaction occurs at the lowest presentation levels of 0 and 8 dB HL. As will be discussed in the following section, this interaction likely occurred since many of the PB-50 words were more often correct at the lower levels than were the W-22 words. Thus, the range of difficulty of the PB-50 words was larger than the range of difficulty of the W-22 words.

Figure 2 presents psychometric functions described by the polynomial equations for the 12 individual subjects (top panels) and for the 100 individual words (middle and bottom panels) for both the W-22 materials (left panels) and the PB-50 materials (right panels). The mean r^2 values for the polynomials describe how well the computed functions agree with the empirical data. For the individual subject data, the mean r^2 values were .99 and .98 for the W-22 and PB-50 materials, respectively. The mean r^2 values for the individual word functions were .98 and .92 for the W-22 and PB-50 materials, respectively, with standard deviations of 0.04 and 0.15. For the W-22 functions, 97 percent of the r^2 values were >.90, whereas with the PB-50 functions, 81 percent of the r^2 values were >.90 and 93 percent were >.75. Thus, the polynomials used to describe the data provided good fits. Displaying functions for 100 words on one panel produced graphic obscurity; thus, the individual word functions are shown by list. The symbols in each panel represent the mean data. The mean presentation levels at which the decade interval percentages

Table 1 Mean Percent Correct Recognition (and SDs) Obtained for Lists 3 and 4 of the W-22 Materials and Lists 8 and 9 of the PB-50 Materials from 12 Listeners with Normal Hearing

dB HL	*W-22*		*PB-50*	
	List 3	*List 4*	*List 8*	*List 9*
0	2.8 (9.7)	2.5 (4.2)	11.8 (11.2)	9.7 (9.7)
8	30.7 (14.1)	33.7 (19.9)	30.0 (22.7)	28.2 (21.8)
16	64.2 (20.2)	61.8 (24.4)	52.0 (27.1)	50.8 (28.0)
24	77.8 (15.4)	80.0 (21.2)	66.2 (25.9)	64.2 (27.7)
32	86.8 (16.0)	88.5 (16.1)	73.2 (25.8)	73.0 (25.7)
40	95.0 (10.2)	90.7 (16.5)	77.5 (25.3)	73.8 (27.4)
48	95.2 (9.1)	94.3 (11.8)	78.0 (26.0)	78.0 (27.4)
56	98.3 (5.6)	98.2 (5.7)	78.7 (27.0)	80.3 (23.1)

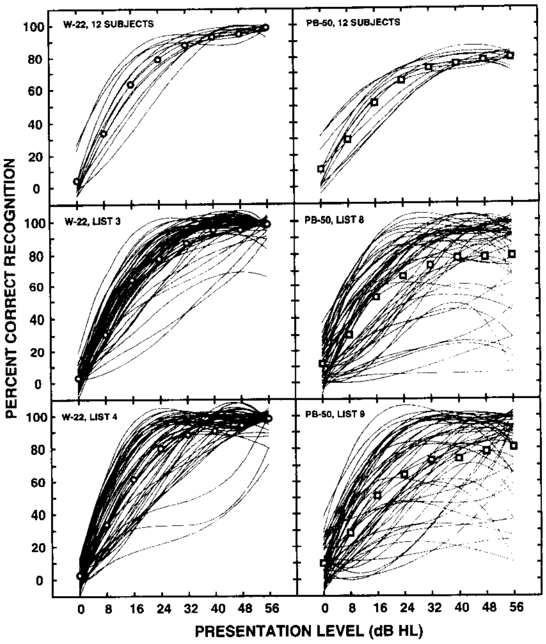

Figure 2 *Top panels:* The mean psychometric functions (best-fit, third-degree polynomials) for each of the 12 subjects for the W-22 materials (lists 3 and 4, left side) and the PB-50 materials (lists 8 and 9, right side). The mean data for the subjects are depicted with open symbols. *Bottom panels:* The mean psychometric functions (best-fit, third-degree polynomials) for each of the words for the 12 subjects. The functions for each list are depicted on separate panels with the mean data shown as open symbols.

correct were obtained and the standard deviations by word (parentheses) and by subject (square brackets) are presented in Table 2. The functions for the majority of words exceeded 80 percent correct at the highest presentation level (56 dB HL). There were, however, exceptions, more so with the PB-50 materials than with the W-22 materials. Twenty-five of the PB-50 functions failed to reach 80 percent correct, with many in the 20 percent to 40 percent correct range regardless of the presentation level (e.g., *bolt, key, nuts, shack, sip,* and *spud*). The polynomial functions for the PB-50 words in this low percent correct range, including those functions that "roll over," accurately reflect the recognition performance obtained by the subjects. Only one of the W-22 words (*owes*) failed to reach 80 percent correct. For these reasons, the data in Table 2 are restricted to 80 percent correct (W-22 lists) and 70 percent correct (PB-50 lists). The spread of the functions in Figure 2 and standard deviations in Table 2 indicates that (1) the intersubject variability is about the same for the two sets of materials (average SDs: 4.8 dB and 5.4 dB for the W-22 and PB-50 lists), (2) the intersubject variability is about half of the interword variability for either set of material (average SDs: W-22 lists, 4.8 dB vs 7.2 dB; PB-50 lists, 5.4 dB vs 11.2 dB), and (3) the interword variability is more for the PB-50 materials (average SDs: 11.2 dB) than for the W-22 materials (average SDs: 7.2 dB). Finally, in Table 2, the mean slopes of the functions for the words at the 50 percent correct point (and standard deviations) are listed. The variability of the slopes of the functions is somewhat more for the PB-50 materials (1.6%/dB) than it is for the W-22 materials (1.2%/dB). From the vertical spread of the individual word functions at the lowest two presentation levels in Figure 2 (0 and 8 dB HL), many of the PB-50 words were more often correct at these lower levels than were the W-22 words. The range of difficulty of the PB-50 words is larger than the range of difficulty of the W-22 words. Thus, both the location and slope characteristics of the PB-50 materials are more variable (less

Table 2 Mean Decibel Hearing Level (ANSI, 1996) at Which the Decade Interval Percent Correct Recognitions Were Obtained for the W-22 and PB-50 Materials

% Correct		*W-22*			*PB-50*	
10	2.5	(2.4)	[2.6]	2.6	(6.9)	[4.0]
20	5.0	(3.6)	[3.3]	7.1	(9.8)	[4.5]
30	7.9	(5.3)	[3.9]	9.4	(9.0)	[5.0]
40	1.4	(8.5)	[4.5]	13.6	(11.9)	[5.4]
50	14.1	(8.3)	[4.9]	17.1	(12.4)	[5.9]
60	17.5	(9.2)	[5.4]	21.2	(13.7)	[6.1]
70	21.3	(10.1)	[7.1]	26.4	(15.0)	[7.1]
80	25.6	(10.4)	[6.7]			
Slope at 50% (%/dB)	3.1	(1.2)	[0.7]	1.9	(1.6)	[0.3]

The standard deviations are shown in parentheses (for the words) and in square brackets (for the subjects). The mean slope (and standard deviation) for the functions at the 50% correct point are listed at the bottom of the table. The data were calculated from the best-fit, third-degree polynomials used to fit the data for each word (n = 100) or each subject (n = 12).

homogeneous) than the same characteristics for the W-22 materials.

As observed in Figure 1, the slopes of the mean functions for the W-22 word lists are steeper than the slopes of the mean functions for the PB-50 word lists. The slopes of the mean functions were calculated using the following three methods: (1) the instantaneous slope computed from the first derivative of the orthogonal polynomial used to fit each set of data, (2) the traditional linear slope that assumes a linear relation between the 20 percent and 80 percent correct points and is calculated simply as $\Delta y / \Delta x$, and (3) the average slope obtained by averaging the instantaneous slopes at 10 percent intervals from the 20 percent to the 70 percent to 80 percent correct points. The instantaneous slopes of the mean W-22 function ranged from 4.3 percent/dB at 20 percent correct to 1.5 percent/dB at 80 percent correct. The 3.1 percent/dB instantaneous slope at 50 percent correct closely mirrored the 3.0 percent/dB average slope but was a little steeper than the 2.8 percent/dB linear slope. Although the PB-50 function did not reach the 80 percent correct point, the relations among the various slopes were similar. The instantaneous slopes ranged from 3.0 percent/dB at 20 percent correct to 0.9 percent/dB at 70 percent correct. The instantaneous slope at the 50 percent point was 1.9 percent/dB, the average slope was 2.2 percent/dB, and the linear slope was 2.3 percent/dB across the 20 percent to 70 percent correct range. The slope data from the mean functions illustrate two relations. First, the slopes at the lower ends of the functions are steeper by a factor greater than two than the slopes at the higher ends of the functions. Second, the instantaneous slope at the 50 percent correct point provides a good approximation either of the average slope or the linear slope of the function over the 20 percent to 70 percent to 80 percent correct points.

The slope characteristics observed with the mean functions (see Fig. 1) were also observed with the functions for the individual words (see Fig. 2). Figure 3 presents a bivariate plot of the slope characteristics of the psychometric functions of the individual 100 W-22 words (top) and 100 PB-50 words (bottom). The values on the abscissa are the linear slopes of the individual word functions calculated between the 20 percent and 70 percent correct points (PB-50 lists) and between the 20 percent and 80 percent correct points (W-22 lists). The values on the ordinate are the instantaneous slopes of the individual word functions calculated from the polynomial equa-

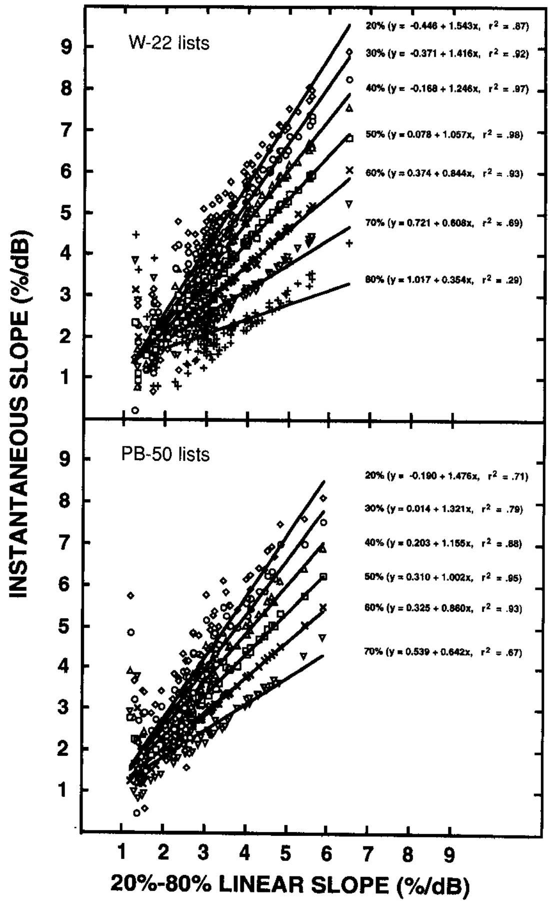

Figure 3 Bivariate plots of the slope characteristics of the psychometric functions of the individual 100 W-22 words and 100 PB-50 words. The linear slopes of the individual word functions between the 20% and 70% to 80% correct points are on the abscissa and the instantaneous slopes of the individual word functions calculated from the polynomial at the decade increments between 20% and 70% to 80% correct are on the ordinate. The equations are the linear regressions fit to each set of 100 datum points.

tions at each of the decade percent correct points between 20 percent and 70 percent to 80 percent correct. The linear regressions used to fit the 100 datum points at each of the percent correct points are listed beside the corresponding data. Many of the 600 to 700 datum points in each panel are superimposed. First, as can be seen from both the graphic and equation representations (b coefficients), the slopes are steeper at the lower percent correct ranges on the functions (the lower presentation levels) than at the higher percent correct ranges (the higher presentation levels). With the W-22 lists, at the 20 percent and 80 percent correct points, the slopes from the linear equations are 1.543 and 0.354, respectively. Likewise, with the PB-50 lists, at the 20 percent and 70 percent correct points, the slopes from the linear equations are 1.476 and 0.642, respectively. Examination of the word recognition functions reported in the literature revealed that most of those functions also were characterized by slopes that were steeper in the lower percent correct recognition ranges (10% to 40%) than in the higher percent correct recognition ranges (60% to 90%) (Tillman and Carhart, 1966; Kopra et al, 1968; Wilson et al, 1976; Beattie and Warren, 1983; Beattie, 1989). Perhaps this relation between the slopes at the opposite ends of the functions suggests listening and response strategies that are different for the various presentation levels. Also from Figure 3, the slopes of the linear regressions for the instantaneous slopes at 50 percent correct and the linear slopes are essentially unity (viz., W-22: 1.057; PB-50: 1.002). This second relation supports findings from the mean data in Figure 1 that the instantaneous slope at the 50 percent correct point is the best predictor of the slope of the function calculated as the linear slope between the 20 percent and the 70 percent to 80 percent correct points. This relation is not surprising as the 50 percent point is in the middle of the 20 percent to 70 percent to 80 percent range.

Finally, it is instructive to examine the data at the individual word and subject level to see how the performances by the subjects contribute to the mean performance on an individual word. Figure 4 presents the individual subject (straight lines) and mean (circles) psychometric functions for two of the W-22 words, *are* (top panel) and *add* (bottom panel). As anticipated, the majority of performances of the subjects incremented from 0 percent to 100 percent correct in one 8-dB step (solid lines). The slopes of these functions are 12.5 percent/dB, which are substantially steeper than the mean functions for the individual subjects. If smaller presentation increments had been used, then steeper functions would have occurred. The exceptions are illustrated by the dashed lines that indicate performance for one subject with each word modulated between 0 percent and 100 percent correct over several presentation levels. These modulating functions probably represent "noise" in the data, which is reflecting some aspect of the attentiveness of the listeners. With *are*, one subject maintained 100 percent correct recognition even at the lowest presentation level. Thus, even at the level of the individual words, the variability of the recognition performances by the subjects is quite varied.

The data from this report demonstrate on subjects with normal hearing that the slope of a mean psychometric function for word recognition is substantially influenced by the location and slope of the psychometric functions of the individual words that contribute to the mean data. The more homogeneous (less variable) the psychometric characteristics of the individual items are, the steeper the slope of the mean psychometric function. One may speculate that the evaluation of word recognition abilities of listeners with sensorineural hearing loss is even more susceptible to the homogeneity factors considered in this report involving listeners with normal hearing. Perhaps this homogeneity/variability factor is one reason that accounts for the word recognition functions for listeners with sensorineural hearing loss as a group being more gradual than functions for the same materials established on listeners with normal hearing. Finally, the current data again urge that caution be used when analyzing and reporting group data, especially when the group is not homogeneous as is the case of any group of individuals with impaired hearing.

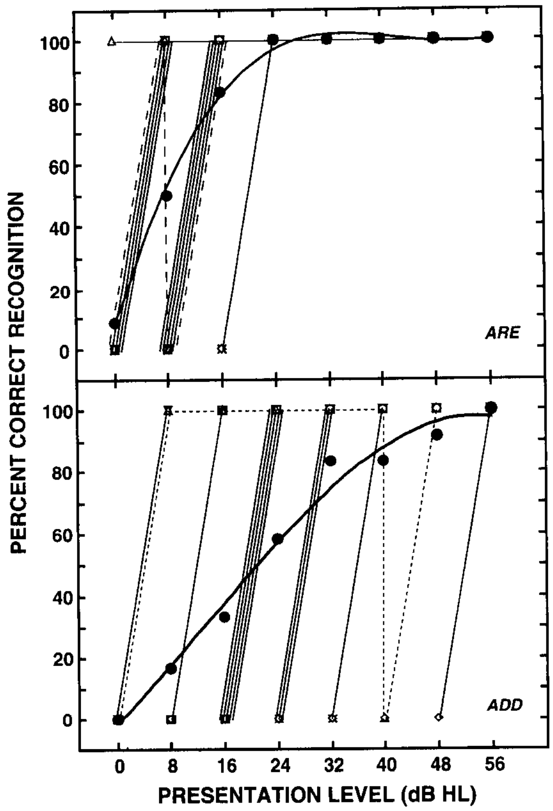

Figure 4 The individual subject (straight lines) and mean (circles) psychometric functions for two of the W-22 words, *are* (top panel) and *add* (bottom panel). The dotted or dashed lines represent the performances of subjects that fluctuated between 0% correct and 100% correct.

Acknowledgments. This project was made possible by funding to the Auditory Research Laboratory at Mountain Home from the Rehabilitation, Research and Development Service and Medical Research Service, Department of Veterans Affairs. The second author is on a Career Development Award sponsored by the Rehabilitation, Research and Development Service. Appreciation is expressed to Lynne Marshall, PhD, who located a copy of the Harris (1948) paper. Portions of this paper were presented at the Annual Convention of the American Academy of Audiology Convention, April 1998, Los Angeles.

REFERENCES

American National Standards Institute. (1996). *American National Standard Specification for Audiometers.* (ANSI S3-1996). New York: ANSI.

Beattie RC. (1989). Word-recognition functions for the CID W-22 test in multitalker noise for normally hearing and hearing-impaired subjects. *J Speech Hear Disord* 54:20–32.

Beattie RC, Warren V. (1983). Slope characteristics of CID W-22 word functions in elderly hearing-impaired listeners. *J Speech Hear Disord* 48:119–127.

Department of Veterans Affairs. (1991). *Speech Recognition and Identification Materials, Disc 1.1.* Long Beach, CA: VA Medical Center.

Harris JD. (1948). *Some Suggestions for Speech Reception Testing. Progress Report No. 2 on BuMed Project NM-003-021.* New London, CT: Naval Medical Research Laboratory, Naval Submarine Base.

Heckendorf AL, Wiley TL, Wilson RH. (1997). Performance norms for the VA compact disc versions of CID W-22 (Hirsh) and PB-50 (Rush Hughes) word lists. *J Am Acad Audiol* 8:163–172.

Hirsh IJ. (1952). *The Measurement of Hearing.* New York: McGraw Hill.

Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. (1952). Development of materials for speech audiometry. *J Speech Hear Disord* 17:321–337.

Kopra LL, Blosser D, Waldron DL. (1968). Comparison of Fairbanks Rhyme Test and CID Auditory Test W-22 in normal and hearing-impaired listeners. *J Speech Hear Res* 11:735–739.

Kruel EJ, Bell DW, Nixon JC. (1969). Factors affecting speech discrimination test difficulty. *J Speech Hear Res* 12:281–287.

Luce PA. (1986). A computational analysis of uniqueness points in auditory word recognition. *Percept Psychophys* 39:155–159.

Luce PA, Pisoni DB. (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear Hear* 19:1–36.

Miller GA, Heise GA, Lichten W. (1951). The intelligibility of speech as a function of the context of the test materials. *J Exp Psychol* 41:329–335.

Rintelmann WF, Schumaier DR, Jetty AJ, Burchfield SA, Beasley DS, Mosher NA, Mosher RA, Penley ED. (1974). Six experiments on speech discrimination utilizing CNC monosyllables. *J Auditory Res* 2:1–30.

Silverman SR, Hirsh IJ. (1955). Problems related to the use of speech in clinical audiometry. *Ann Otol Rhinol Laryngol* 64:1234–1244.

Tillman TW, Carhart R. (1966). *An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words.* Northwestern University Auditory Test No. 6. Brooks Air Force Base, TX: USAF School of Aerospace Medicine.

Wilson RH, Coley KE, Haenel J, Browning K. (1976). Northwestern University Auditory Test No. 6: normative and comparative intelligibility functions. *J Am Audiol Soc* 1:221–228.

Wilson RH, Margolis RH. (1983). Measurements of auditory thresholds for speech stimuli. In: Konkle DF, Rintelmann WF, eds. *Principles of Speech Audiometry.* Baltimore: University Park, 79–126.

Wilson RH, Zizz CA, Shanks JE, Causey GD. (1990). Normative data in quiet, broadband noise, and competing message for Northwestern University Auditory Test No. 6 by a female speaker. *J Speech Hear Disord* 55:771–778.