J Am Acad Audiol 17:385–397 (2006)

Word Recognition of Digit Triplets and Monosyllabic Words in Multitalker Babble by Listeners with Sensorineural Hearing Loss

Richard H. Wilson*†
Christopher A. Burks*†
Deborah G. Weakley*†

Abstract

In an initial experiment (Wilson and Weakley, 2004), word recognition was assessed with six digit triplets presented at 14 signal-to-babble ratios (S/B) in 2 dB steps. An abbreviated version of the protocol was developed for clinic use involving three digit triplets at 7 S/Bs in 4 dB steps. The purpose of this experiment was to examine the relationship between the two digit protocols with comparisons made with other variables including age, pure-tone thresholds, subjective measures of understanding speech in quiet and in noise, and word recognition of monosyllabic words in quiet and in babble. Ninety-six listeners with sensorineural hearing loss participated. For equivalent performance, the short version of the digit triplets required (1) a 2.6 dB more favorable S/B than the long version and (2) a 15.1 dB less favorable S/B than the words. Age, hearing loss, and subjective evaluation of the ability to understand speech in quiet and in noise were not related to performance on digits or words in multitalker babble.

Key Words: Auditory perception, digit triplets, hearing loss, speech perception, word recognition in multitalker babble

Abbreviations: NU No. 6 = Northwestern University Auditory Test No. 6; S/B = signal-to-babble ratio; S/N = signal-to-noise ratio

Sumario

En un experimento inicial (Wilson y Weakley, 2004) se evaluó el reconocimiento de palabras con la presentación de tripletas de seis dígitos a 14 tasas diferentes de señal/balbuceo (S/B) in pasos de 2 dB. Se desarrolló una versión resumida del protocolo para uso clínico, utilizando tripletas de tres dígitos a 7 tasas S/B en pasos de 4 dB. El propósito de este experimento fue examinar la relación entre los protocolos de dos dígitos, realizando comparaciones con otras variables que incluían edad, umbrales de tonos puros, medidas objetivas de comprensión de lenguaje en silencio y en ruido, y de reconocimiento de palabras monosilábicas en silencio y en balbuceo. Participaron noventa y seis sujetos con hipoacusia sensorineural. Ante rendimientos equivalentes, la versión corta de tripletas de dígitos requirió (1) un S/B más favorable de 2.6 dB que la versión larga y (2) un S/B menos favorable de 15.1 dB que con

*James H. Quillen VA Medical Center, Mountain Home, Tennessee; †Departments of Surgery and Communicative Disorders, East Tennessee State University, Johnson City, Tennessee

Richard H. Wilson, Ph.D., VA Medical Center, Audiology (126), Mountain Home, TN 37684; Phone: 423-979-3561; Fax: 423-979-3403; E-mail: RICHARD.WILSON2@MED.VA.GOV

palabras. La edad, la hipoacusia y la evaluación subjetiva de la capacidad de entender lenguaje en silencio y en ruido, no se relacionó con el rendimiento con dígitos o palabras en medio de balbuceo de hablantes múltiples.

Palabras Clave: Percepción auditiva, tripletas de dígitos, hipoacusia, percepción del lenguaje, reconocimiento de palabras en balbuceo de hablantes múltiples

Abreviaturas: NU No. 6 = Prueba auditiva No. 6 de la Universidad Northwestern; S/B = tasa de señal/balbuceo; S/N = tasa de señal/ruido

Most people with hearing loss complain most often that they can hear speech but have difficulty understanding speech, especially in background noise. Data from numerous studies substantiate this complaint (Hirsh, 1950; Carhart and Tillman, 1970; Keith and Talis, 1970; Tillman et al, 1970; Olsen et al, 1975; Dubno et al, 1984; Gordon-Salant, 1987; Beattie, 1989; Pekkarinen et al, 1990; Souza and Turner, 1994; Divenyi and Haupt, 1997a, 1997b, 1997c; Wiley et al, 1998; Wilson et al, 2003). The ability of patients to understand speech in background noise typically is not evaluated in audiology practice (Martin et al, 1998). Sentence materials involving presentations in background noise (e.g., the Synthetic Sentence Identification materials [SSI; Speaks and Jerger, 1965], the Speech Perception In Noise [SPIN; Kalikow et al, 1977; Bilger et al, 1984], the Connected Speech Test [CST; Cox et al, 1987; Cox et al, 1988], the Speech-In-Noise test [SIN; Killion and Villchur, 1993], and the Hearing In Noise Test [HINT; Nilsson et al, 1994]) have been developed but have not gained widespread use, probably because audiologists are accustomed to evaluating word-recognition abilities with monosyllabic words in quiet. For this reason to gain clinical acceptance, our efforts have been directed at the development of a word-recognition paradigm in background noise. Over the past several years, through a series of studies (Wilson et al, 1996; Wilson and Strouse, 2002; Wilson, 2003; Wilson and Weakley, 2004), experimental versions of a monosyllabic word test in multitalker babble and a digit triplet test in multitalker babble were developed. Wilson and Weakley (2004) observed for both word and digit materials that (1) listeners with hearing loss required a 6–12 dB better signal-to-babble ratio (S/B) to maintain the same level of recognition performance obtained by listeners with normal hearing, and (2) for both groups of listeners, equivalent recognition performances were obtained when the words were at a 15 dB more favorable signal-to-babble ratio than the digits. The relation of importance was that both materials provided bimodal distributions that separated the two groups of listeners.

In our initial exploratory study with digit triplets, which was precipitated by the Smits et al (2004) study involving digit triplets in background noise to screen hearing over the telephone, the protocol was lengthy and involved the presentation of six digit triplets at each of 14 presentation levels from 6 to -20 dB S/B in 2 dB steps (Wilson and Weakley, 2004). With the ultimate goal of application in the clinic or as a hearing-screening instrument, it was obvious that the protocol had to be shortened. Based on the previous data from 24 listeners with normal hearing and 48 listeners with high-frequency sensorineural hearing loss, two parameters of the protocol were changed. First, the 2 dB step size was increased to 4 dB steps from 4 to -20 dB S/B. Second, the number of digit triplets at each level was reduced from six triplets (18 digits) in the original study to three triplets (nine digits). These two abbreviations reduced the time required to administer the protocol by about 75%. The purpose of this experiment was to examine the shortened protocol primarily in terms of two trials and two randomizations. As a comparative reference, performance on monosyllabic words in multitalker babble was evaluated in the protocol. In addition to ear, trial, and randomization effects on recognition performance of the digits-in-babble, the design permitted examination of the relationships among other variables including age, pure-tone thresholds, subjective measures of understanding speech in quiet and in noise, and word recognition in quiet.

METHODS

Materials

Development of the digit triplets and monosyllabic words in multitalker babble are described in detail in Wilson and Weakley (2004) and Wilson (2003), respectively. Briefly, the digits 1 through 10 (excluding 7, which is bisyllabic) were time-locked with unique segments of multitalker babble with 300 msec before and after each digit. The amplitude of each of the nine digits then was adjusted to produce the signal-to-babble ratios from 4 to -20 dB in 4 dB steps; the level of the babble was fixed. At each of the seven signal-to-babble ratios (63 digits) and in quiet (9 digits), three digits triplets for each of the eight conditions randomly were selected. The triplets were formed by concatenating the required digit/babble segments with 300 msec (25 msec rise/fall times) added before the first digit of the triplet and after the last digit of the triplet. Following each digit triplet, a 4 sec interstimulus interval (quiet) was inserted. Each randomization of the 72 digits was approximately three minutes.

As with the digits, the monosyllabic words from Northwestern University Auditory Test No. 6 (NU No. 6; Tillman and Carhart, 1966; Department of Veterans Affairs, 1998) were paired with unique segments of multitalker babble. Then based on psychometric data from both listeners with normal hearing and listeners with high-frequency sensorineural hearing loss, psychometric data were used to determine which ten words would be presented at each of seven signal-to-babble ratios (70 total words) (Wilson, 2003). The babble, which was fixed in level, was presented continuously during the presentation of the 70 words with a 2.7 sec interval between words for a total time of five minutes. For the words in quiet condition, two 25-word lists from NU No. 6, List 1, were compiled, each of which was two minutes. All materials were recorded on an audio compact disc (Hewlett-Packard, Model DVD200i).

Subjects

The 96 patients (mean = 63.8 years) were recruited consecutively from the ongoing Audiology Clinics at Mountain Home. The inclusion criteria were that thresholds at 500 and 1000 Hz were ≤30 dB HL (American National Standards Institute, 1996) and ≤40 dB HL, respectively, with thresholds above 1000 Hz <100 dB HL. Word recognition in quiet was not an inclusion criterion. The means and standard deviations for the pure-tone thresholds and word-recognition abilities by ear are listed in Table 1. The 4.0–5.5 dB differences between ears at 2000–4000 Hz with lower thresholds for the left ear were statistically significant, but for practical purposes the differences were not considered remarkable. Additionally, the audiologic data from all listeners were consistent with sensorineural hearing loss. Human subjects approval was obtained, and the experimental procedures followed the standards of the institutional review board. Subjects who made a special trip to the laboratory were compensated.

Procedures

Initially, the listeners were asked to rate on a scale of 1 (no difficulty) to 10 (extreme difficulty) how much difficulty they had understanding (1) speech in quiet, and (2)

Table 1. Measures of Central Tendency for the Age, Pure-Tone Audiogram, and Percent Correct Word Recognition (WR) for the 96 Listeners

	AGE (Years)	Frequency in Hertz 250	500	1000	2000	3000	4000	8000	WR (% Correct)
Left Ear									
Mean	63.8	18.5	19.3	23.2	47.2	64.3	70.5	65.4	78.3
SD	9.1	7.6	6.4	8.9	17.4	15.2	14.6	16.3	16.9
Right Ear									
Mean		19.9	19.5	23.5	43.2	58.8	66.4	65.4	79.6
SD		8.5	7.1	8.3	18.6	17.2	17.2	15.8	16.4
LE – RE		-1.4	-0.2	-0.3	4.0	5.5	4.2	0.0	-1.3
t-test (p)		0.031	0.568	0.583	0.006	0.001	0.015	0.972	0.360

Note: Data for the left and right ears are listed.

speech in noise (Wilson et al, 2005a, appendix). Second, the digit triplets were presented with 24 listeners assigned to each of the following four presentation combinations: R2L3, R3L2, L2R3, and L3R2. In this counterbalanced scheme, each of the variables was given an equal number of times in each position. That is, 48 listeners received the stimuli to the right ear (R) first followed by the stimuli to the left ear (L); the remaining 48 listeners received the stimuli to the ears in the opposite order. Likewise, Randomizations 2 and 3 each were presented an equal number of times on Trial 1 and Trial 2. The digits were presented monaurally in quiet at 80 dB SPL and in multitalker babble at signal-to-babble ratios ranging in 4 dB decrements from 4 to -20 dB (84 to 60 dB SPL) with the babble fixed at 80 dB SPL. Each of the nine digits was presented at each signal-to-babble ratio. The test was terminated when all words at one level were incorrect.

Third, the monosyllabic words were presented binaurally in multitalker babble at 24 to 0 dB S/Bs (104 to 80 dB SPL) in 4 dB decrements with the babble again fixed at 80 dB SPL. Ten words were presented at each of the seven levels. Again, the test terminated when all words were incorrect at one level. Finally, List 1 of the NU No. 6 was administered binaurally with the first 25 words presented at 80 dB SPL and the second 25 words presented at 104 dB SPL.

All testing was conducted in a double-wall sound booth with the signals reproduced on a CD player (Sony, Model CDP-497), fed through an audiometer (Grason-Stadler, Model 61), and delivered to TDH-50P earphones encased in Telephonics P/N 510C017-1 cushions. Data collection was completed in about 20 minutes with the verbal responses of the listeners recorded into a spreadsheet.

RESULTS

The Spearman-Kärber equation (Finney, 1952) was used to calculate the 50% points on the data for each subject from the three babble conditions (two monaural digit conditions and binaural words). The data in Table 2 provide a brief overview of the mean and standard deviation data (from the two subjective questions, the digits-in-babble, the words-in-babble, and the words-in-quiet) that will be discussed throughout this section. The individual digit data are plotted in Figure 1. The three-panel figure shows the digit data by ear (top), by randomization (middle), and by trial (bottom). The diagonal line in each panel represents equal performance with the numbers above, on, and below the line indicating the number of respective data points. The filled symbols in each panel are the mean data. The shaded region in the lower left corner of each panel represents the 90th percentile of performance at the 50% point (-10.9 dB S/B) on the digit triplets by 16 listeners with normal hearing (Wilson et al, 2005a).

First, consider the data by ear in the top panel of Figure 1. Data points above the diagonal line represent better performance on the digit triplets presented to the right ear (51) than on the digit triplets presented to the left ear (41). The difference between means was 0.4 dB (Table 2). Second, there was no effect of randomization of the digit presentations (middle panel) with equal number of points above and below the diagonal line (46). Third from the bottom panel of Figure 1, 63 listeners had better performance on the second trial than on the first trial, whereas 29 listeners had better performance on the first trial than on the second trial. These graphic analyses were confirmed by numeric analysis using a two-between (ear, randomization), one-within

Table 2. Measures of Central Tendency for the Subjective Questions, the 50% Points (dB S/B), and the Percent Correct Word Recognition Obtained in Quiet

	Question		50% Point in dB S/B			Words in Quiet (%)	
				Digits			
	Quiet	Noise	Words	Left	Right	60 dB HL	84 dB HL
Mean	2.9	6.1	12.3	-3.2	-3.6	85.3	94.2
SD	1.5	1.8	3.9	3.1	3.1	16.2	9.0
Max	8.0	10.0	23.6	4.7	6.0	100.0	100.0
Min	1.0	2.0	3.6	-9.1	-10.0	36.0	60.0
Range	7.0	8.0	20.0	13.8	16.0	64.0	40.0

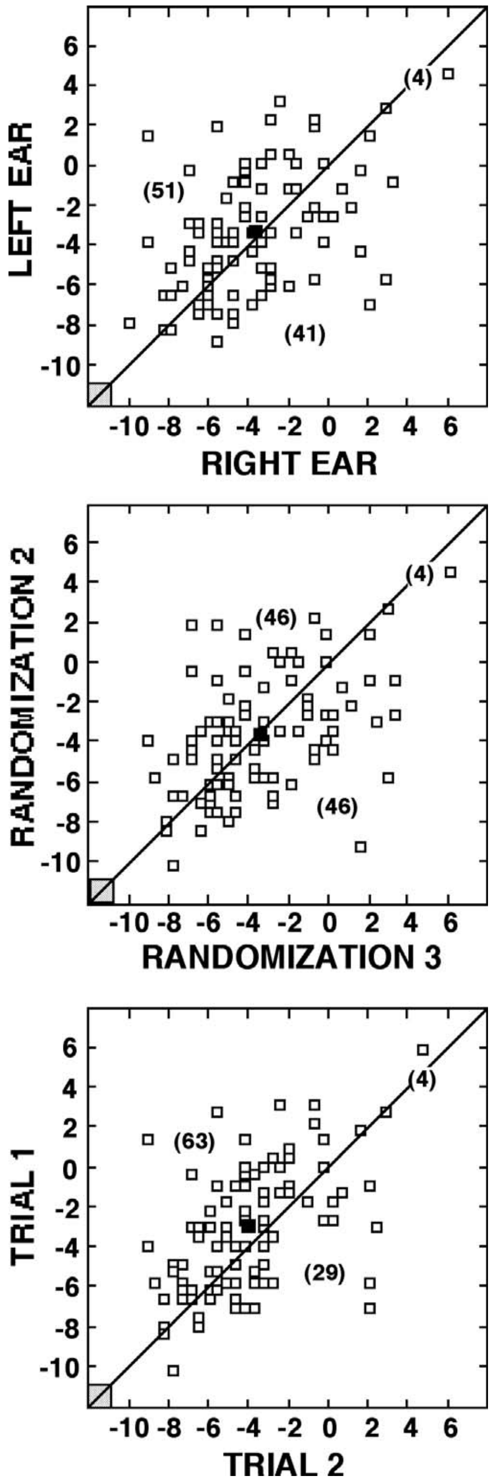

Figure 1. Bivariate plots of the digit data by ear (top), by randomization (middle), and by trial (bottom). The diagonal line in each panel represents equal performance with the numbers above, on, and below the line indicating the number of respective data points. The filled symbols in each panel are the mean data. The shaded region in the lower left corner of each panel represents the 90th percentile of performance at the 50% point (-10.9 dB S/B) on the digit triplets by 16 listeners with normal hearing (Wilson et al, 2005a).

(trial) mixed model analysis-of-variance (ANOVA). The main effects of ear and randomization were not significantly different. The ANOVA indicated that the trial main effect of the recognition performance was significantly better on the Trial 2 than performance on Trial 1 ($F[1,92] = 14.72$, $p = .0002$); the 0.7 dB difference, however, was not of practical importance.

In Figure 2, the mean correct recognition data are plotted as a function of the respective presentation levels of the digits. The lines through the datum points are the best-fit, third-degree polynomials used to describe the data. The decibel values on each set of

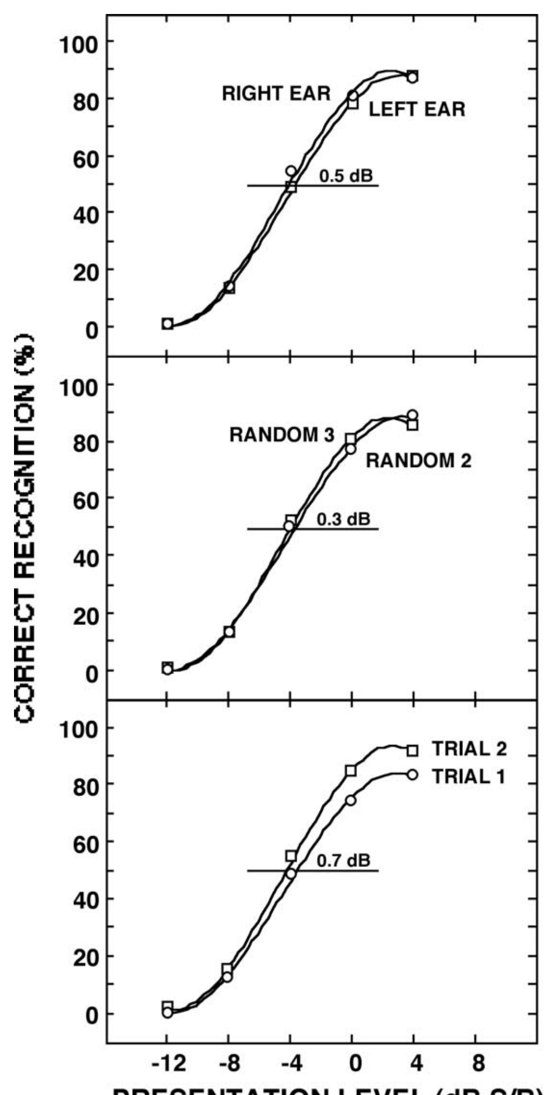

Figure 2. The mean correct recognition data for the digits in babble are plotted as a function of the respective presentation levels. The lines through the datum points are the best-fit, third-degree polynomials used to describe the data. The decibel values on each set of data indicate the difference between the two functions in each panel at the 50% points.

data indicate the difference between the two functions in each panel at the 50% points. Whereas the data in Figure 1 were restricted to recognition performances at the 50% points, the data in Figure 2 demonstrate that throughout the range of recognition performances, the respective pairs of variables had almost identical characteristics. Based on the findings reported in Figures 1 and 2, the data for the digit triplets were combined and are used throughout the remainder of this paper.

Figure 3 is a bivariate plot of the 50% recognition performances for the words in multitalker babble from the individual listeners (abscissa) versus the 50% point from the digits in babble (ordinate). The shaded region on the figure represents the 90th percentile of normal performance (6.0 dB S/B for words [Wilson et al, 2003] and -10.9 dB S/B for digit triplets [Wilson et al, 2005a]). Again, the thresholds were established on the individual data with the Spearman-Kärber equation. The mean of the performance on the words in babble (Table 2) was 12.3 dB with a standard deviation of 3.9 dB. The obvious feature of the data in Figure 3 is that only three of the datum points are within the 90th percentile for performance by listeners with normal hearing. Interestingly, these three datum points are in the normal range only for the words in babble, not for the digits in babble. The gradual slope of the regression function (0.4 dB/dB) indicates a general direct relationship between performances on the word and digit materials, but the variability is substantial.

The mean psychometric functions for both the digit (squares) and word (circles) materials in babble are illustrated in Figure 4. The vertical lines through the datum points depict the standard deviations, and the curvilinear lines are the best-fit, third-degree polynomials used to describe the data. The thin function to the left represents the digit data from Wilson and Weakley (2004) at the levels that correspond to the levels used in the current study. For equal performance throughout the range of recognition performances, the words required an ~15 dB more favorable signal-to-babble ratio than was required by the digit triplets, which is close to the differences between performances on digits and words reported previously by Miller et al (1951) and Wilson and Weakley (2004). (The differences among the functions in Figure 4 are considered in detail in the "Discussion" section.) The slopes of the

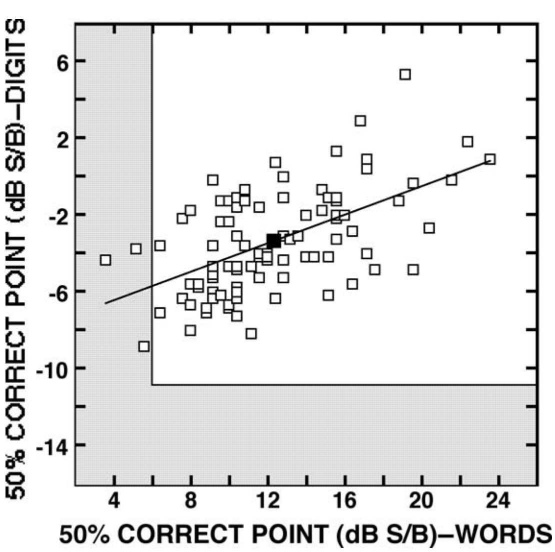

Figure 3. A bivariate plot depicting the 50% recognition performances for the words in multitalker babble from the individual listeners (abscissa) versus the 50% point from the digits in babble (ordinate). The shaded region on the figure represents the 90th percentile of normal performance (6.0 dB S/B for words [Wilson et al, 2003] and -10.9 dB S/B for digit triplets [Wilson et al, 2005a]).

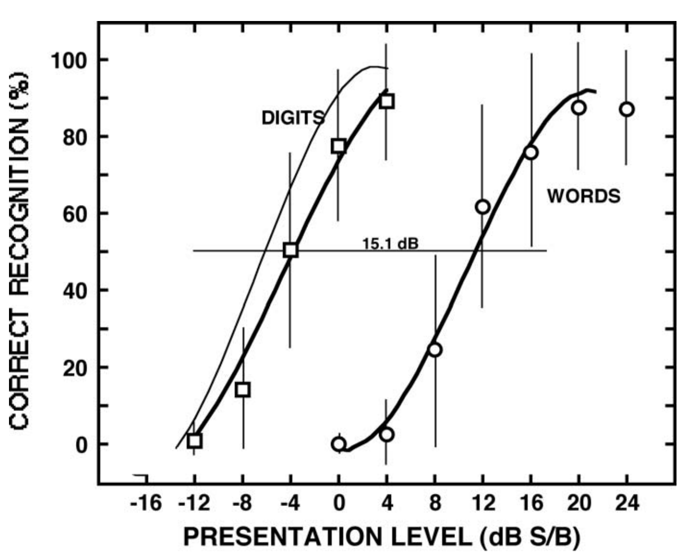

Figure 4. The mean psychometric functions for the digit and word materials in babble are shown. The vertical lines through the datum points depict the standard deviations, and the curvilinear lines are the best-fit, third-degree polynomials used to describe the data. The thin line (left) is the function for digit performance obtained in the Wilson and Weakley (2004) study.

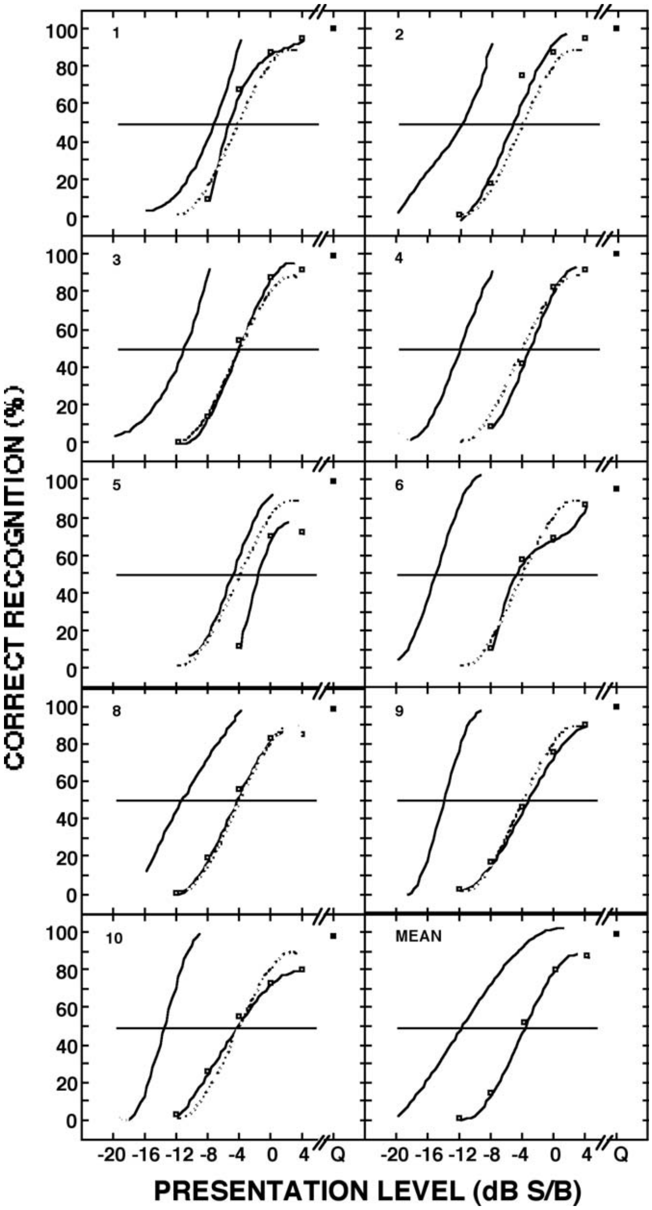

Figure 5. The psychometric functions for the individual digits in multitalker babble and in quiet (Q). For comparison in each panel, the psychometric function for the corresponding digit obtained in our laboratory from 36 listeners with normal hearing is shown (thin lines) along with the mean function for the digits (dashed lines).

functions at the 50% points were the same, 6.7%/dB. As expected with a group of listeners with hearing loss, the variability is substantial, especially on the dynamic portions of the functions.

Figure 5 displays the mean psychometric functions for the individual digits in multitalker babble and in quiet (Q). For comparison in each panel, the psychometric function for the corresponding digit obtained by McArdle et al (2005) from 36 listeners with normal hearing is shown (thin lines) along with the mean function for the digits obtained in the current study (dashed lines). Again, the line through the datum points is the best-fit, third-degree polynomial used to describe the data. The 50% points calculated from the individual data and the slopes at the 50% points are presented in Table 3 along with similar data from 36 subjects with normal hearing (McArdle et al, 2005). The obvious features in the graph and table are the large interdigit variabilities both in terms of the location of the functions in the Cartesian coordinates and the slopes of the functions. The data from both groups of listeners indicate that the digit "5" had the highest 50% correct point, which is consistent with earlier data (Wilson and Weakley, 2004). The variability of the 50% points was greater for the listeners with normal hearing than for the listeners with hearing loss. This difference is in large part owing to the 50% point for "five" that was an outlier in the data for the listeners with normal hearing. The differences between the 50% points for the listeners with normal hearing and with hearing loss ranged from 1.9 dB ("one") to 10.7 dB ("nine") with a mean difference of 7.2 dB. The slopes of the digit functions for the listeners with hearing loss ranged 14.4%/dB from 6.4%/dB ("10") to 20.8%/dB ("five"), which is a larger range than the 8.1%/dB range for the listeners with normal hearing (7.5 to 15.6%/dB).

The data in Figure 6 are the mean pure-tone thresholds (500, 1000, 2000, and 4000 Hz) plotted as a function of the signal-to-babble ratio (dB) at which 50% correct recognition was obtained (ordinate) for the monosyllabic words (top panel) and digit triplets (bottom panel). The line through the datum points is a linear regression fit to the data. Again, there is a shotgun pattern of the datum points for both the words and digit triplets. The slopes of the linear regressions are basically flat (digits, 0.016 %/dB; words, -0.025%/dB) indicating that for the subjects included in this study, hearing loss and performance on the digits and on the words varied independently. The same analysis was applied to the digit data and the individual pure-tone thresholds at 500, 1000, 2000, and 4000 Hz; the slopes of the regressions for the respective frequencies

Table 3. 50% Points (dB S/B) and Slopes at the 50% Points Calculated from the Polynomials Used to Fit the Data for 36 Listeners with Normal Hearing (McArdle et al, 2005) and for the 96 Listeners (from Figure 5)

	Normal Hearing McArdle et al (2005)		Hearing Loss Current Study		Difference
No.	50% Point	Slope	50% Point	Slope	at 50%
1	-7.5	10.8	-5.6	13.0	1.9
2	-12.0	7.6	-5.5	10.0	6.5
3	-11.3	10.5	-4.3	10.5	7.0
4	-12.2	11.4	-3.3	10.1	8.9
5	-5.2	11.5	-1.6	20.8	3.6
6	-15.4	13.1	-5.1	8.4	10.3
8	-11.6	7.5	-4.6	8.9	7.0
9	-14.2	15.6	-3.5	7.9	10.7
10	-13.9	15.5	-4.6	6.4	9.3
Mean of Functions	-11.5	11.5	-4.2	10.7	7.2
Standard Deviations	3.3	2.9	1.3	4.2	3.0
Mean Function	-12.0	6.5	-4.0	9.1	8.0

Note: The differences between the 50% points (dB) also are shown. The mean of the functions is the mean of the individual digit functions whereas the mean function is the function used to describe the mean data.

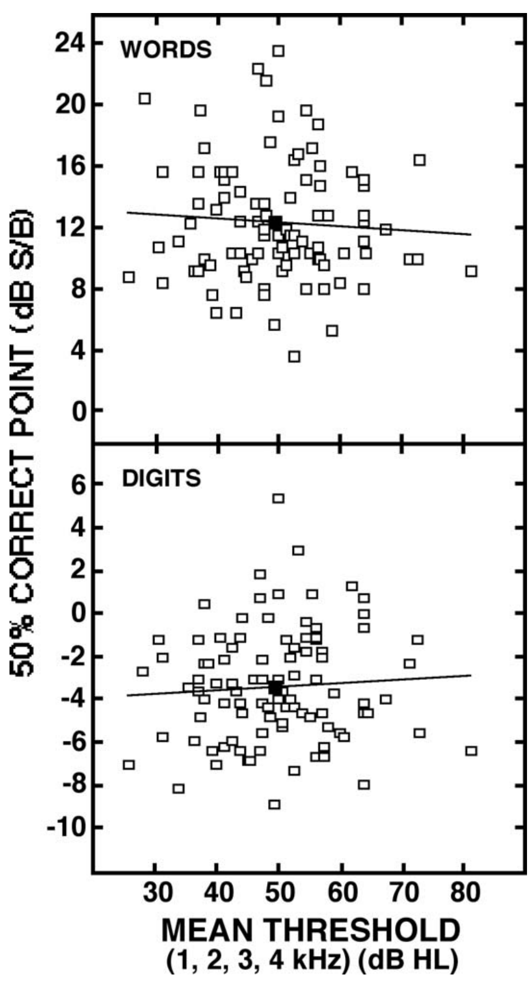

Figure 6. Bivariate plots of the mean pure-tone thresholds (500, 1000, 2000, and 4000 Hz) (abscissa) plotted as a function of the signal-to-babble ratio (dB) at which 50% correct recognition was obtained (ordinate) for the monosyllabic words (top) and digit triplets (bottom).

were –0.044, -0.002, 0.049, and 0.058 dB/dB, which confirms that the graphic analysis of the mean pure-tone data accurately reflect the components of the whole.

The data in Figure 7 are a series of bivariate plots featuring age, pure-tone threshold averages, and the signal-to-babble ratio (dB) at which 50% correct recognition was obtained. The pattern of data in each plot may be characterized as "shotgun," meaning that there is little or no organization to the pattern. In the top panel, age (abscissa) is plotted against the pure-tone average at 500, 1000, 2000, and 4000 Hz (ordinate). The linear regression fit to the data has a slightly positive slope (0.21 dB/year), suggesting little change in hearing as a function of age, at least for the individuals included in this study. The regression lines in both the middle and lower panels are practically flat (-0.04 and 0.01 dB/year, respectively), which indicates no relationship between age and recognition performance on either the digits or the words.

The individual data from the subjective noise question are plotted in Figure 8 (words top panel and digits bottom panel) with the rating for the noise (1–10) on the abscissa and recognition performance in multitalker babble plotted on the ordinate. Again, "shotgun" best describes the pattern of the bivariate plots. Both regression lines have slopes that approach zero, -0.1 points/dB (words) and -0.2 points/dB (digits). Thus, as earlier reports consistently have demonstrated, the relationship between the perceived difficulty that a listener has in background noise and the objective measure of the difficulty that a

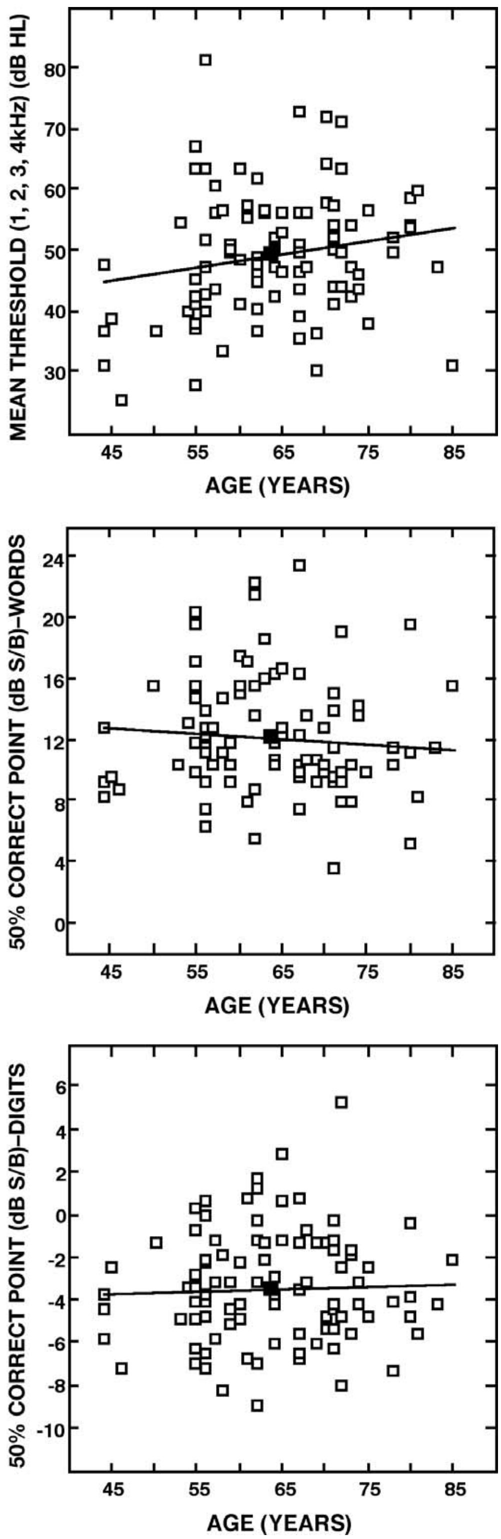

Figure 7. Bivariate plots of the age of the listeners (abscissa) and mean pure-tone thresholds (top), the signal-to-babble ratio at which 50% correct recognition was obtained for the digits (middle), and the signal-to-babble ratio at which 50% correct recognition was obtained for the monosyllabic words (bottom).

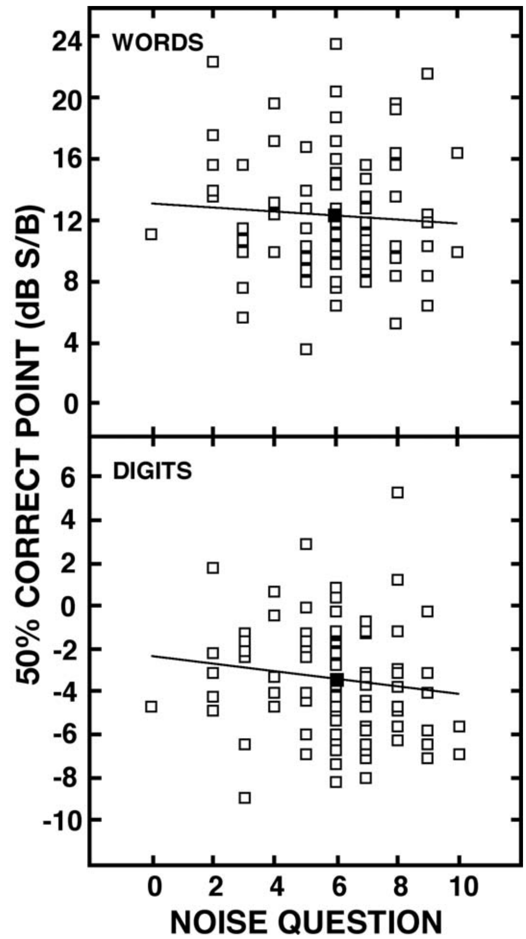

Figure 8. Bivariate plots of the 1 to 10 rating obtained on the noise question (abscissa) and the signal-to-babble ratio at which the 50% correct recognition point (ordinate) was obtained for the monosyllabic words (top) and digit triplets (bottom).

listener has in background noise is basically nonexistent (Rowland et al, 1985; Wilson et al, 2005b). Although not shown, the relationship between the responses to the subjective quiet and noise questions was examined. All of the listeners indicated about three times more difficulty listening in noise than in quiet (see Table 2). In a bivariate plot of the data, the slope of the regression line was 0.4 point/point meaning that as the difficulty understanding speech in noise increased 1 point, the difficulty in quiet increased 0.4 points. Interestingly, the slope would have been steeper except there were several listeners who had 1 and 2 points in quiet and 8 or 9 points in noise. No subject reported listening in quiet easier than listening in noise.

The lack of relationships observed in the bivariate plots in Figures 7 and 8 was confirmed by a multiple linear regression analysis in which age, pure-tone octave thresholds from 250–8000 Hz, subjective ratings of speech understanding in quiet and noise, and word-recognition ability in quiet were entered as independent variables to examine digit recognition performance in noise (Jerger, 1963). None of the independent variables were found to be significant predictors of digits recognition in noise (F = 1.74; $p > .05$). The adjusted r^2 for this model was 14.9%, and none of the standardized coefficients were found to be significant.

Figure 9 depicts the word-recognition performance in quiet (ordinate) versus the performance in multitalker babble (abscissa) for the 60 and 84 dB HL presentation levels in quiet. The shaded region indicates the area of normal performance for the two variables. At both levels, only 3 of the 96 listeners performed in the normal range for listening in quiet and in babble. As one would expect, recognition performance in quiet improved as the presentation level increased. The majority of the listeners had word recognition in quiet that was in the "normal" range, that is, 80% correct or better. At 60 dB HL there were 24 listeners with performances <80% correct, with only 6 listeners <80% correct at 84 dB HL. The majority of the listeners had word recognition in quiet that was in the "normal" range, that is, 80% correct or better.

DISCUSSION

Two versus Three Digit Sets

The major question posed in this study was this: How does recognition performance on the abbreviated digit triplet protocol compare with the performance on the more extensive protocol used in the original Wilson and Weakley (2004) study? Comparison of the two digit functions in Figure 4 (function with the squares and the thin line from the original study) provides insight. The level for the 50% point on the original digit function was -6.3 dB S/B whereas the level for the 50% point on the current digit function was -3.7 dB S/B. The 2.6 dB difference can be attributed either to a group effect (i.e., the two groups are simply different) or to a "practice" effect. The listeners in the original protocol had substantially more exposure to the listening/response task than did the listeners in the current study. Recall that the original paradigm involved six digit triplets presented at each level and 2 dB steps whereas the current protocol involved three digit triplets presented at each level and 4 dB steps. Support for a practice effect is found in the data in Figure 2 (bottom panel) from the current study. The data in Figure 2 indicate that at the 50% point, 0.7 dB better performance was obtained on Trial 2 than on Trial 1, which was a difference that was accentuated at the higher signal-to-babble ratios. Trial 1 provided the practice on the listening task that produced the better performance on Trial 2. The relations between the functions in Figure 2 and Figure

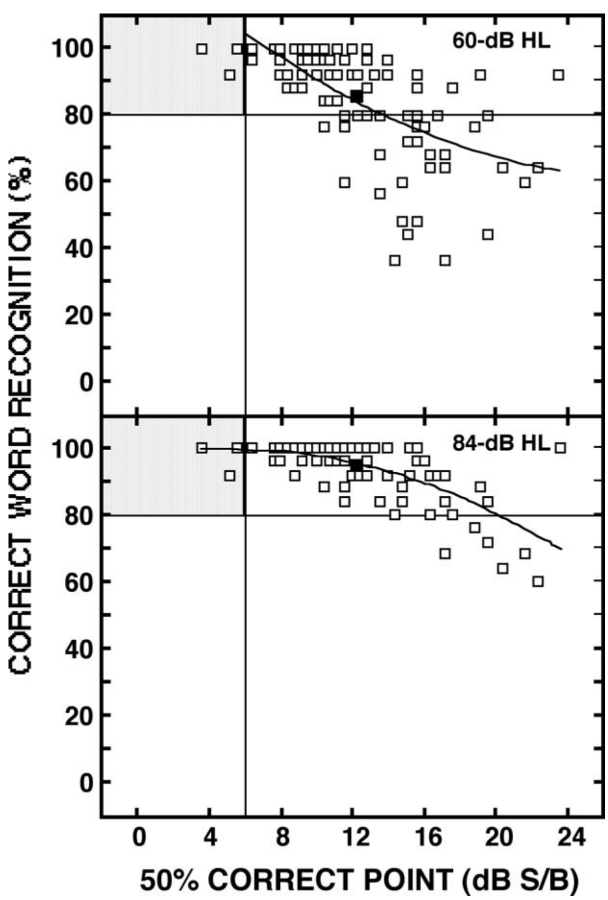

Figure 9. Bivariate plots of the word-recognition performance in quiet (ordinate) versus the signal-to-noise ratio at which the 50% point was obtained in multitalker babble (abscissa) for the 60 (top) and 84 dB HL (bottom) presentation levels in quiet. The shaded region indicates the area of normal performance for the two variables.

Table 4. Measures of Central Tendency for the 50% Points (dB S/B) Obtained from the 96 Listeners

	Ear		Trial		Randomization	
	Left	Right	1st	2nd	2	3
Mean	-3.2	-3.6	-2.8	-4.0	-3.4	-3.5
SD	3.1	3.1	3.2	2.9	3.0	3.2
Max	4.7	6.0	6.0	4.7	4.7	6.0
Min	-9.1	-10.0	-10.0	-9.1	-10.0	-9.1
Range	13.8	16.0	16.0	13.8	14.7	15.1
50% Point	-3.7	-4.2	-3.6	-4.3	-3.8	-4.1
Slope	8.8	9.3	8.5	9.7	8.8	9.5

Note: Also included are the 50% points (dB S/B) and slopes at the 50% points (%/dB) computed from the third-degree polynomials used to describe the mean data in Figure 2.

4 closely resemble one another. The 50% point on the word function in the Wilson and Weakley (2004) study (11.2 dB S/B) was identical to the 50% point on the word function from the current study (11.4 dB S/B). Thus, on the words-in-babble task, the listeners in the original and current studies had the same recognition performance. Although digits- and words-in-babble have ranges of recognition performances that are distinctly different in the Cartesian coordinate system, collectively, the Wilson and Weakley (2004) study and the current study demonstrate that the relation of importance is that digits and monosyllabic words provide the same differences in performances between listeners with normal hearing and listeners with hearing loss.

The data in Figures 1 and 2 and Tables 2 and 4 demonstrated that the digits triplets produced equal performances for the right and left ear and for the two randomizations used in the protocol. As mentioned above, there was a 0.7 dB mean improvement between Trial 1 and Trial 2 that was statistically significant. When an instrument like digit-triplet recognition is used clinically, a <1 dB difference between trials must be considered uneventful, especially with reference to overall performance. The one thing that the 0.7 dB difference does suggest is that clinic protocol should involve testing the suspected better ear (e.g., pure-tone sensitivity) first, just as is done with most audiologic protocols.

The mean psychometric functions plotted in Figure 4 for the digit and word materials demonstrated a consistent difference at the various levels of recognition performance. At the 50% point on the functions there was a 15.1 dB difference with the monosyllabic words requiring a more favorable signal-to-babble ratio. In the previous study (Wilson and Weakley, 2004), an 18 dB difference was observed between digits and monosyllabic words whereas Miller et al (1951, Table 1) made a similar observation with a 19 dB difference between the signal-to-noise ratios (S/N) of the two types of materials (digit threshold, -14 dB S/N; monosyllabic word threshold, 5 dB S/N). Although this difference in recognition performances between the two materials is widely quoted, there are aspects of the relation that are suspect, especially when considered in terms of the relation between detection and recognition. Typically, the difference between a detection threshold and a recognition threshold for speech materials is on the order of 7–9 dB with recognition requiring a more favorable signal-to-noise ratio than detection (Hawkins and Stevens, 1950). There is no reason to believe that the detection thresholds for monosyllabic digits like those used in this study should be any different from the detection thresholds for other monosyllabic words.

Digits versus Words

If detection for digits and monosyllabic words were the same, then for equal recognition performance why do monosyllabic words require a 15–19 dB more favorable signal-to-noise ratio than digits? Here, focus must be shifted to the response task involved with the two materials, that is, close-set and open-set response tasks. The response task involved with the digits is a closed set in that there are only nine alternative responses possible (digits 1–10, excluding 7). The number of alternative responses involved with the monosyllabic words is so large that the response task can be considered an open set. Thus, the difference between performances on digits and words is not a difference between the materials, but rather, it is a difference between open- and closed-set response paradigms. In an earlier experiment when the same monosyllabic words were used in both an open-set response paradigm (oral recall) and a closed-set response paradigm (in which the subject responds by pointing to written words or pictures in a quadrant arrangement, i.e., four possible responses), equal performance was obtained when the words in the open-set paradigm were presented at levels 9–12 dB higher than the levels for the same words in the closed-set paradigm (Wilson and Antablin, 1982). Intuitively, increasing the alternatives in the closed set from four to nine (the number of digits in the current study) would reduce the 9–12 dB difference by some unspecified amount. Using this line of reasoning, the difference between recognition performances on the digits and words in the current study and in the Miller et al study should be <9 dB.

Other factors are thought to have contributed to the differences observed in recognition of digits and monosyllabic words by Miller et al (1951) and Wilson and Weakley (2004). Miller et al used altered pronunciations of some of the digits, for example, “niner” for “nine.” An important consideration of the Miller et al study was that the two listeners involved in the study alternated between speaking the stimuli via monitored live voice and listening to the stimuli through a 200–3000 Hz band-pass system. Conceivably, these factors produced a listening task that was more akin to digit detection than to digit recognition. To explain from personal experience, when a listener becomes extremely familiar with a speech stimulus, the recognition of that stimulus is based on listening cues like the recognition of microstructures of the word and rhythm of the word, not the traditional factors associated with a recognition task. With this altered type of recognition task, the relation between detection and recognition performances can be diminished from the traditional 7–9 dB to 2–3 dB. Applying this to the Miller et al study, it is probable that the digit recognition function was only a couple of decibels above the theorized digit detection function. At best this would account for 5–6 dB of the 19 dB difference between the performances on the digits and words. Additionally, the Wilson and Weakley (2004) study attributed 3–5 dB of the difference between the recognition performances on digits and monosyllabic words to calibration differences. The digits were calibrated to 0 vu whereas the carrier phrases with the monosyllabic words were calibrated to 0 vu, not the words that were always several units below 0 vu. With this 3–5 dB correction, the difference between the digits and words in the Wilson and Weakley (2004) study would be on the order of 13 dB. Thus, different reasons are thought to contribute to the differences in recognition performances on digits and on monosyllabic words in the Miller et al and Wilson and Weakley (2004) studies. To resolve this issue, an experiment involving detection and recognition tasks is underway in which the same speaker speaks both stimuli and the same calibration rules are applied to both stimuli. With the stimulus variables held constant, then the only difference between performances on the two materials should be at most 8–9 dB that is attributable to the issue of closed-set versus open-set response paradigms.

Acknowledgment. The Rehabilitation Research and Development Service, Department of Veterans Affairs, supported this work through a Merit Review, the Auditory and Vestibular Dysfunction Research Enhancement Award Program (REAP), and a Senior Research Career Scientist award to the first author.

REFERENCES

American National Standards Institute. (1996) *Specification for Audiometers* (ANSI S3.6-1996). New York: American National Standards Institute.

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.

Bilger RC, Nuetzel JM, Rabinowitz WM, Rzeczkowski C. (1984) Standardization of a test of speech perception in noise. *J Speech Hear Res* 27:32–48.

Carhart R, Tillman TW. (1970) Interaction of competing speech signals with hearing loss. *Arch Otolaryngol* 91:273–279.

Cox RM, Alexander GC, Gilmore C. (1987) Development of the connected speech test (CST). *Ear Hear* 8:119S–126S.

Cox RM, Alexander GC, Gilmore C, Pusakulich KM. (1988) Use of the connected speech test (CST) with hearing-impaired listeners. *Ear Hear* 9:198–207.

Department of Veterans Affairs. (1998) *Speech Recognition and Identification Materials, Disc 2.0.* Mountain Home, TN: VA Medical Center.

Divenyi PL, Haupt KM. (1997a) Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. I. Age and lateral asymmetry effects. *Ear Hear* 18:42–61.

Divenyi PL, Haupt KM. (1997b) Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. II. Correlation analysis. *Ear Hear* 18:100–113.

Divenyi PL, Haupt KM. (1997c) Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation. *Ear Hear* 18:189–201.

Dubno JR, Dirks DD, Morgan DE. (1984) Effects of age and mild hearing loss on speech recognition in noise. *J Acoust Soc Am* 76:87–96.

Finney DJ. (1952) *Statistical Method in Biological Assay.* London: C. Griffen.

Gordon-Salant S. (1987) Age-related differences in speech recognition performance as a function of test format and paradigm. *Ear Hear* 8:277–282.

Hawkins JE, Stevens SS. (1950) The masking of pure tones and of speech by white noise. *J Acoust Soc Am* 22:6–13.

Hirsh IJ. (1950) Binaural hearing aids: a review of some experiments. *J Speech Hear Disord* 15:114–123.

Jerger J. (1963) Viewpoint. *J Speech Hear Res* 6: 203–206.

Kalikow DN, Stevens KN, Elliot LL. (1977) Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am* 61:1337–1351.

Keith RW, Talis HP. (1970) The use of speech in noise in diagnostic audiometry. *J Aud Res* 10:201–204.

Killion MC, Villchur E. (1993) Kessler was right—partly: but SIN test shows some aids improve hearing in noise. *Hear J* 46:31–35.

Martin FN, Champlin CA, Chambers JA. (1998) Seventh survey of audiometric practices in the United States. *J Am Acad Audiol* 9:95–104.

McArdle RA, Wilson RH, Burks CA. (2005) Speech recognition in multitalker babble using digits, words, and sentences. *J Am Acad Audiol* 16(9):726–739.

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.

Nilsson M, Soli SD, Sullivan JA. (1994). Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am* 95:1085–1099.

Olsen WO, Noffsinger D, Kurdziel S. (1975) Speech discrimination in quiet and in white noise by patients with peripheral and central lesions. *Acta Otolaryngol* 80:375–382.

Pekkarinen E, Salmivalli A, Suonpää J. (1990) Effect of noise on word discrimination by subjects with impaired hearing, compared with those with normal hearing. *Scand Audiol* 19:31–36.

Rowland JP, Dirks DD, Dubno JR, Bell TS. (1985) Comparison of speech recognition-in-noise and subjective communication assessment. *Ear Hear* 6:291–296.

Smits C, Kapteyn TS, Houtgast T. (2004) Development and validation of an automatic SRT screening test by telephone. *Int J Audiol* 43:15–28.

Souza PE, Turner CW. (1994) Masking of speech in young and elderly listeners with hearing loss. *J Speech Hear Res* 37:655–661.

Speaks C, Jerger J. (1965) Performance-intensity characteristics of synthetic sentences. *J Speech Hear Res* 9:305–312.

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

Tillman TW, Carhart R, Olsen WO. (1970) Hearing aid efficiency in a competing speech situation. *J Speech Hear Res* 13:789–811.

Wiley TL, Cruickshanks KJ, Nondahl DM, Tweed TS, Klein R, Klein BEK. (1998) Aging and word recognition in competing message. *J Am Acad Audiol* 9:191–198.

Wilson RH. (2003) Development of a speech in multitalker babble paradigm to assess word-recognition performance. *J Am Acad Audiol* 14:453–470.

Wilson RH, Abrams HB, Pillion AL. (2003) A word-recognition task in multitalker babble using a descending presentation mode from 24-dB S/B to 0-dB S/B. *J Rehabil Res Dev* 40:321–328. http://www.vard.org/jour/03/40/4/pdf/Wilson-B.pdf.

Wilson RH, Antablin JK. (1982) The Picture Identification Task, A reply to Dillon. *J Speech Hear Disord* 47:111–112.

Wilson RH, Burks CA, Weakley DG. (2005a) A comparison of word-recognition abilities assessed with digit pairs and digit triplets in multitalker babble. *J Rehabil Res Dev* 42:499–510.

Wilson RH, Burks CA, Weakley DG. (2005b) Word recognition in multitalker babble measured with two psychophysical methods. *J Am Acad Audiol* 16:627–636.

Wilson RH, Oyler AL, Sumrall R. (1996) Psychometric functions for Northwestern University Auditory Test No. 6 in quiet and multitalker babble. Paper presented at the 8th Annual American Academy of Audiology Convention, Salt Lake City.

Wilson RH, Strouse A. (2002) Northwestern University Auditory Test No. 6 in multitalker babble: a preliminary report. *J Rehabil Res Dev* 39:105–113. www.vard.org/jour/02/39/1/wilson.htm.

Wilson RH, Weakley DG. (2004) The use of digit triplets to evaluate word-recognition abilities in multitalker babble. *Semin Hear* 25:93–111.