J Am Acad Audiol 19:548–556 (2008)

A Comparison of Two Word-Recognition Tasks in Multitalker Babble: Speech Recognition in Noise Test (SPRINT) and Words-in-Noise Test (WIN)

DOI: 10.3766/jaaa.19.7.4

Richard H. Wilson*
Wendy B. Cates*

Abstract

Background: The Speech Recognition in Noise Test (SPRINT) is a word-recognition instrument that presents the 200 Northwestern University Auditory Test No. 6 (NU-6) words binaurally at 50 dB HL in a multitalker babble at a 9 dB signal-to-noise ratio (S/N) (Cord et al, 1992). The SPRINT was developed by and used by the Army as a more valid predictor of communication abilities (than pure-tone thresholds or word-recognition in quiet) for issues involving fitness for duty from a hearing perspective of Army personnel. The Words-in-Noise test (WIN) is a slightly different word-recognition task in a fixed level multitalker babble with 10 NU-6 words presented at each of 7 S/N from 24 to 0 dB S/N in 4 dB decrements (Wilson, 2003; Wilson and McArdle, 2007). For the two instruments, both the babble and the speakers of the words are different. The SPRINT uses all 200 NU-6 words, whereas the WIN uses a maximum of 70 words.

Purpose: The purpose was to compare recognition performances by 24 young listeners with normal hearing and 48 older listeners with sensorineural hearing on the SPRINT and WIN protocols.

Research Design: A quasi-experimental, mixed model design was used.

Study Sample: The 24 young listeners with normal hearing (19 to 29 years, mean = 23.3 years) were from the local university and had normal hearing (≤20 dB HL; American National Standards Institute, 2004) at the 250–8000 Hz octave intervals. The 48 older listeners with sensorineural hearing loss (60 to 82 years, mean = 69.9 years) had the following inclusion criteria: (1) a threshold at 500 Hz between 15 and 30 dB HL, (2) a threshold at 1000 Hz between 20 and 40 dB HL, (3) a three-frequency pure-tone average (500, 1000, and 2000 Hz) of ≤40 dB HL, (4) word-recognition scores in quiet ≥40%, and (5) no history of middle ear or retrocochlear pathology as determined by an audiologic evaluation.

Data Collection and Analysis: The speech materials were presented bilaterally in the following order: (1) the SPRINT at 50 dB HL, (2) two half lists of NU-6 words in quiet at 60 dB HL and 80 dB HL, and (3) the two 35-word lists of the WIN materials with the multitalker babble fixed at 60 dB HL. Data collection occurred during a 40–60 minute session. Recognition performances on each stimulus word were analyzed.

Results: The listeners with normal hearing obtained 92.5% correct on the SPRINT with a 50% point on the WIN of 2.7 dB S/N. The listeners with hearing loss obtained 65.3% correct on the SPRINT and a WIN 50% point at 12.0 dB S/N. The SPRINT and WIN were significantly correlated ($r = -0.81$, $p < .01$), indicating that the SPRINT had good concurrent validity. The high-frequency, pure-tone average (1000, 2000, 4000 Hz) had higher correlations with the SPRINT, WIN, and NU-6 in quiet than did the traditional three-frequency pure-tone average (500, 1000, 2000 Hz).

Conclusions: Graphically and numerically the SPRINT and WIN were highly related, which is indicative of good concurrent validity of the SPRINT.

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

---

*VA Medical Center, Mountain Home, TN, and Departments of Surgery and Communicative Disorders, East Tennessee State University, Johnson City, TN

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@VA.GOV

This work was presented at AudiologyNOW! 2008 in Charlotte, NC.

Abbreviations: HFPTA = high-frequency, pure-tone average; MOS = Military Occupational Specialty; MMRB = Medical MOS Retention Board; NU-6 = Northwestern University Auditory Test No. 6; S/N = signal-to-noise ratio; SPRINT = Speech Recognition in Noise Test; VA = Veterans Affairs; WIN = Words-in-Noise Test

Sumario

Antecedentes: La Prueba de Reconocimiento del Lenguaje en Ruido (SPRINT) es un instrumento de reconocimiento de palabras que presenta binauralmente las 200 palabras de la Prueba Auditiva No. 6 de la Universidad Northwestern a 50 dB HL en balbuceo de hablantes múltiples, a una tasa señal-ruido (S/N) de 9 dB (Cord y col., 1992). El SPRINT fue desarrollado y usado por el Ejército como un elemento con mayor validez de predicción de las habilidades de comunicación (que los umbrales de tonos puros o el reconocimiento de palabras en silencio) con asuntos relacionados con la aptitud para el servicio militar desde la perspectiva auditiva del personal del ejército. La Prueba de Palabras en Ruido (WIN) es una tarea ligeramente diferente de reconocimiento de palabras a un nivel fijo de balbuceo de hablantes múltiples con 10 palabras NU-6 presentadas a cada uno de 7 S/N desde 24 a 0 dB S/N en decrementos de 4 dB (Wilson, 2003; Wilson y McArdle, 2007). Para los dos instrumentos, tanto el balbuceo como los hablantes de las palabras son diferentes. El SPRINT usa todas las 200 palabras del NU-6, mientras que el WIN usa un máximo de 70 palabras.

Propósito: El propósito fue comparar el desempeño en reconocimiento de 24 sujetos con audición normal y 48 sujetos mayores con alteración sensorineural, en relación con los protocolos SPRINT y WIN.

Diseño de Investigación: Se utilizó un diseño mixto de modelo, cuasi- experimental.

Muestra del Estudio: Los 24 sujetos jóvenes con audición normal (19 a 29 años, media = 23.3 años) pertenecían a la universidad local y tenían umbrales normales (≤ 20 dB HL; Instituto Nacional Americano de Estándares, 2004) en los intervalos de octava de 250–8000 Hz. Los 48 sujetos mayores con hipoacusia sensorineural (60 a 82 años, media = 69.9 años) tenían los siguientes criterios de inclusión: (1) un umbral en 500 Hz entre 15 y 30 dB HL; (2) un umbral a 1000 Hz entre 20 y 40 db HL, (3) un promedio tonal puro en tres frecuencias (500, 1000 y 2000 Hz) de ≤40 dB HL, (4) puntajes de reconocimiento de palabras en silencio ≥40%, y (5) sin historia de patología en el oído medio o retrococlear, determinado por una evaluación audiológica.

Recolección y Análisis de Datos: Los materiales de lenguaje fueron presentados bilateralmente en el siguiente orden: (1) el SPRINT a 50 dB HL, (2) dos medias listas de las palabras NU-6 en silencio a 60 dB HL y 80 dB HL, y (3) las dos listas de 35 palabras de los materiales del WIN con el balbuceo de hablantes múltiples a un nivel fijo de 60 dB HL. La recolección de datos ocurrió durante una sesión de 40–60 minutos. Se analizó el desempeño en el reconocimiento en cada palabra estímulo.

Resultados: Los sujetos con audición normal obtuvieron un 92.5 % de respuestas correctas en el SPRINT con un punto 50% en el WIN de S/N de 2.7 dB. Los sujetos con hipoacusia obtuvieron un 65.3% de respuestas correcta en el SPRINT y un punto 50% en el WIN a una S/N de 12.0 dB HL. El SPRINT y el WIN correlacionaron significativamente ($r = -0.81$, $p < 0.01$), indicando que el SPRINT tuvo una buena validez concurrente. El promedio tonal a alta frecuencia (1000, 2000, 4000 Hz) tuvo correlaciones mayores con el SPRINT, WIN y NU-6 en silencio que lo que tuvo el promedio tonal puro tradicional en tres frecuencias (500, 1000 y 2000 Hz).

Conclusiones: Gráfica y numéricamente, el SPRINT y el WIN estuvieron altamente relacionados, lo que es indicativo de una buena validez concurrente del SPRINT.

Palabras Clave: Percepción auditiva, pérdida auditiva, percepción del lenguaje, reconocimiento del lenguaje en balbuceo de hablantes múltiples

Abreviaturas: HFPTA = promedio tonal puro de altas frecuencias; MOS = Especialidad Ocupacional Militar; MMRB = Consejo de Retención del MOS Médico; NU-6 = Prueba Auditiva No. 6 de la Universidad Northwestern; S/N = tasa señal-ruido; SPRINT = Prueba de Reconocimiento del Lenguaje en Ruido; VA = Asuntos de Veteranos; WIN = Prueba de Palabras en Ruido

The Army uses a physical profile system known as PUHLES to categorize the general physical condition, upper extremities, hearing, lower extremities, vision (eye), and mental status of soldiers. Within each category a four-point scale is incorporated with 1 being the best rating and 4 being the worst rating. Accordingly, the Army classifies hearing loss by pure-tone thresholds using H-1 to H-4 profiles

Table 1. Decibel Hearing Levels Used to Define the H-1 and H-2 Profiles

	Frequency in Hz				
Profile	500	1000	2000	3 Freq. PTA	4000
H-1 (both ears)	30	30	30	25	45
H-2 (both ears)	35	35	35	30	55
or					
H-2 (better ear)	30	25	25		35

Note: For the first two profiles, the unaided three-frequency pure-tone averages (500, 1000, and 2000 Hz) for each ear are the first determiner with no single threshold being greater than the limits indicated. For the third profile the better ear has to have the indicated minimum thresholds and the poorer ear "may be deaf" (Department of the Army, 2007, table 7-1).

(Department of the Army, 2007). For the H-1 and H-2 profiles, hearing loss (unaided) in either ear can have no pure-tone threshold greater than the limits indicated in Table 1 for the respective categories. If a threshold exceeds the limits at one frequency, then the next higher profile is applied. The H-3 profile, which involves any pure-tone threshold that exceeds the limits set for the H-2 profile, requires an audiological evaluation and for fitness of duty may involve a Medical Military Occupational Specialty (MOS) Retention Board, or MMRB, who determines if the soldier should be (1) retained in his or her current MOS, (2) reclassified into another MOS, or (3) discharged from the Army. The H-4 profile indicates that hearing, even with a hearing aid(s), is sufficiently impaired to preclude continued service in the Army.

Despite its reliance on pure-tone thresholds to profile hearing, both Army audiologists and members of MMRB realized that neither pure-tone thresholds nor speech-recognition abilities assessed in quiet were accurate indicators of the ability of a soldier to communicate in realistic environments involving background noise. Consequently, the Speech Recognition in Noise Test (SPRINT) was developed as a more valid predictor of communication abilities than were pure-tone thresholds or word recognition in quiet, that is, fitness for duty from a hearing perspective (Cord et al, 1992). The SPRINT scores, in conjunction with other medical, occupational, and administrative considerations, are used by the MMRB to make the most appropriate determination regarding retention, reclassification, or discharge of the soldier. Administration of the SPRINT is required by Army regulations for soldiers having H-3 profiles whose medical records are subject to review by an MMRB. The Army administers thousands of SPRINTs annually. Additionally, in the current era of Iraqi Freedom and Enduring Freedom, many private audiologists and VA (Veterans Affairs) audiologists are being requested to administer the SPRINT on military personnel being activated and inactivated.

The SPRINT involves presentation of monosyllabic words in background noise simultaneously to both ears. Specifically, the 200 monosyllabic words that constitute the Northwestern University Auditory Test. No. 6 (NU-6; Tillman and Carhart, 1966) are presented in a multitalker babble at one (9 dB) signal-to-noise ratio (S/N). In keeping with the purpose for which the test is administered, the large number of words is included to achieve highly stable scores, following the binomial model (Thornton and Raffin, 1978). The 9 dB S/N was intended to produce 95–100% performances in young adults with normal hearing, but widely varying performances among soldiers having the H-3 profiles. During development, the SPRINT was administered to 319 soldiers with H-3 profiles (Cord et al, 1992). Mean recognition performance was 81.5% correct (163 of 200 words) with a range from 32.5% (65 words correct) to 98% (196 words correct). Test-retest recognition performances, which were established on 30 soldiers, were 82.3 and 83.4%, respectively, and were highly correlated ($r = 0.93$, $p < .01$).

Despite being developed for the specific purpose described above, use of the SPRINT has been proposed for other applications within the government, involving different clinical populations. Consequently, and because of its widespread use within the military, the purpose of the current study was to extend the available data on the SPRINT to include young adults with normal hearing and older listeners with hearing loss who were >60 years of age, which is above the age range of the 319 soldiers on whom the SPRINT originally was evaluated. A second purpose was to examine the concurrent validity of the SPRINT by comparing the recognition performances obtained on the SPRINT by the younger and older listeners with performances obtained on the Words-in-Noise test (WIN) that also uses the NU-6 materials in a multitalker babble paradigm but at multiple S/Ns (Wilson, 2003; Wilson and Burks, 2005; Wilson and McArdle, 2007). The concurrent validity of the WIN was established in previous studies (McArdle et al, 2005; Wilson and McArdle, 2005; Wilson et al, 2007) that examined the relationship between performances on the WIN and other speech-recognition paradigms including monosyllabic words in quiet, the HINT (Hearing In Noise Test; Nilsson et al, 1994), the BKB-SIN™ (Niquette et al, 2003), and the Quick-SIN™ (Quick Speech-in-Noise test; Killion et al, 2004).

METHODS

Materials

The NU-6 words (Randomization C) used in the SPRINT were recorded by Auditec of St. Louis (www.auditec.com) (male speaker) and are mixed with a six-talker, multitalker babble at 9 dB S/N. Because 200 words were administered, the test required about 20 min.

The WIN consists of 70 NU-6 words spoken by the VA female speaker (Department of Veterans Affairs, 2006). The WIN test has the following characteristics (Wilson, 2003): (1) 70 monosyllabic words from the same recorded version of NU-6 used for the recognition measures in quiet (Department of Veterans Affairs), which enables the evaluation of recognition performance in quiet and in multitalker babble with the same materials spoken by the same speaker; (2) ten unique words presented at each of seven signal-to-babble ratios in 4 dB decrements from 24 to 0 dB S/N; (3) each word is time-locked to a unique segment of multitalker babble, which reduces variability; (4) the level of the babble, which is presented continuously, is fixed and the level of the words varied; (5) the interval between words is 2.7 sec; (6) the 50% point is quantifiable with the Spearman-Kärber equation (Finney, 1952); and (7) a stopping rule that terminates the test sequence when the ten words at one level are incorrect. For the current study, the 70-word WIN list was divided into two 35-word lists that were administered sequentially, and the results from the two runs were combined (Wilson and Burks, 2005).

Half lists from List 4 of the NU-6 materials were used to evaluate word recognition in quiet with the half lists counterbalanced between two presentation levels. The two levels, 60 and 80 dB HL, correspond to the presentation level of the words in the 0 and 20 dB S/N conditions of the WIN. Data from these two levels help ensure that audibility was not an issue when interpreting the WIN data, especially at the less favorable S/Ns. All of the speech stimuli were calibrated to a common 1000 Hz tone using a vu meter (Tascam, Model MU-40) and were recorded on a CD (Hewlett Packard, Model DVD200i).

Subjects

Twenty-four young listeners between the ages of 19 years and 29 years (mean = 23.3 years, SD = 2.4) with normal hearing and 48 older listeners between the ages of 60 years and 82 years (mean = 69.9 years, SD = 6.7) with sensorineural hearing loss participated in the study. The young listeners were recruited from the local university and had normal hearing (≤20 dB HL; American National Standards Institute, 2004) at the 250–8000 Hz octave intervals. The inclusion criteria for the listeners with hearing loss were as follows: (1) age between 60 and 85 years, (2) a threshold at 500 Hz between 15 and 30 dB HL, (3) a threshold at 1000 Hz between 20 and 40 dB HL, (4) a three-frequency pure-tone average (500, 1000, and 2000 Hz) of ≤40 dB HL, (5) word-recognition scores in quiet ≥40%, and (6) no history of middle ear or retrocochlear pathology as determined by an audiologic evaluation. The pure-tone criteria applied to the ear with the better thresholds. Pure-tone symmetry was not considered, as word-recognition performance in noise is essentially the same (±1 dB) for monaural and binaural conditions when listening under earphones (Holma et al, 1997; Wilson, 2003). The average difference between the pure-tone thresholds for the right and left ears of the listeners with hearing loss was <1 dB. The average audiogram and standard deviations for the right ear are shown in Figure 1.

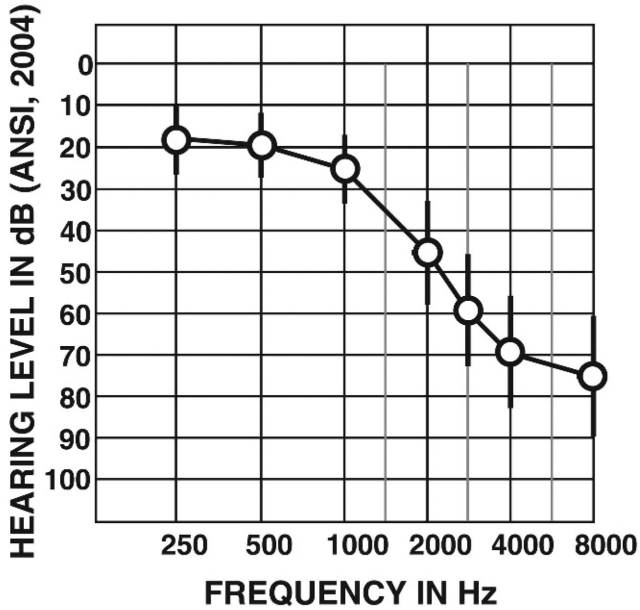

Figure 1. The mean audiogram for the right ears of the 48 listeners with hearing loss is shown along with the standard deviations represented by the vertical lines.

Procedures

Data collection occurred during a 40–60 minute session. Following the informed-consent process, pure-tone thresholds were obtained. The speech materials were presented bilaterally in the following order: (1) SPRINT protocol administered at 50 dB HL, which is the presentation level used by the Army, (2) two half lists of NU-6 words were administered in quiet with the first half list at 60 dB HL and the second half list at 80 dB HL, and (3) the two 35-word lists of the WIN materials were administered with the multitalker babble fixed at 60 dB HL. For the SPRINT the instructions given in the Cord et al (1992) paper were used. The speech materials were reproduced on a CD player (Sony, Model CDP-497) that was routed through an audiometer (GSI, Model 61). All stimuli were presented under TDH-50P earphones while the participants were seated in a double-walled sound booth (IAC, Model 1204). With the word stimuli, the

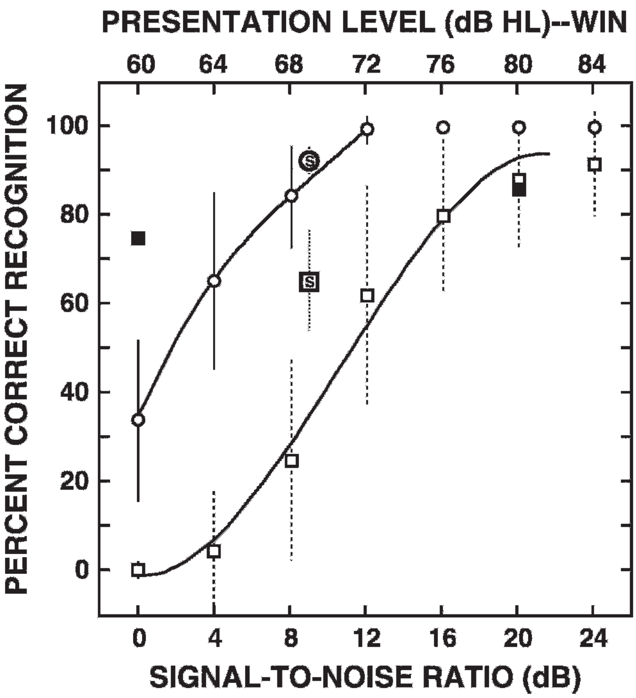

Figure 2. The mean datum points obtained from the 24 listeners with normal hearing (circles) and the 48 listeners with hearing loss (squares) for the SPRINT (large symbols with "S" included), for the WIN (small open symbols), and for the NU-6 presented in quiet (filled squares). The vertical lines through the datum points represent ±1 standard deviation. The lines through the datum points are the best-fit, third-degree polynomials used to describe the data.

verbal responses of the listeners were recorded into a spreadsheet.

RESULTS AND DISCUSSION

The overall results from the 24 listeners with normal hearing (circles) and from the 48 listeners with sensorineural hearing loss (squares) are shown in Figure 2. The two mean data points for the SPRINT are depicted by the larger symbols denoted by "S" at the 9 dB S/N. The mean WIN datum points are depicted with the smaller symbols connected with third-degree polynomials used to describe the data. The vertical lines through the various datum points indicate ±1 standard deviation.

Consider first the SPRINT data in Figure 2. The mean for the listeners with normal hearing was 92.5% correct (SD = 2.4%), whereas the mean for the older listeners with hearing loss was 65.3% (SD = 11.2%). The 92.5% for the listeners with normal hearing was slightly below the performance level targeted by Cord et al (1992) of 95–100% with only five of the listeners performing in the targeted range. Performance by the listeners with hearing loss was on average 27.2% below the mean performance by the listeners with normal hearing and 16.2% below the mean performance (81.5%) reported by Cord et al for the 319 soldiers with H-3 profiles. This lower performance by the listeners with hearing loss in this study was expected as they were older than the Cord et al sample and they had poorer pure-tone thresholds than the Cord et al soldiers, especially in the 2000–3000 Hz range.

Although the SPRINT is scored as a composite of 200 words, the performances on each of the four NU-6 lists were available. For the listeners with normal hearing the scores for Lists 1-4 were 89.0, 93.7, 92.8, and 94.4% with standard deviations ranging from 3.0 to 4.4%. The respective scores for the listeners with hearing loss were: 62.1, 66.8, 67.3, and 65.2% with standard deviations ranging from 11.6 to 13.1%. Although no statistical analyses were conducted on these data, it is interesting that performance by both groups of listeners on List 1 was about 5% (~1 dB, assuming a slope of 5%/dB) poorer than the performances on the other lists. Probably this relation is reflecting a "learning curve or practice effect" by the listeners that occurred during List 1, which was always given first. It is well established that practice effects improve performance, especially in difficult listening conditions such as noise at less than optimal S/Ns (e.g., Wilson, Bell, et al, 2003; Sweetow and Sabes, 2006). List 1 enabled the listeners to improve their listening skills in the multitalker babble background noise that was reflected by the improved performances on the remaining lists. Alternatively, the difference in performances between List 1 and the remaining lists could be attributed to interlist differences. This reason for the differences probably can be discounted as the Auditec recordings of the NU-6 materials have been shown to have good inter-list equivalency (Wilson et al, 1976).

The WIN data in Figure 2 illustrate the typical relation between recognition performances on the WIN by listeners with normal hearing and listeners with hearing loss. From the functions in the figure, the 50% points were 1.8 and 11.3 dB S/N, respectively, for the two groups of listeners, which is a 9.5% difference. The slopes of the respective functions at the 50% points were 8.1 and 6.7%/dB, again an expected relationship with the function for the listeners with hearing loss being more gradual owing to their increased variability that is demonstrated by the standard deviations (vertical lines through the datum points in the figure). The filled squares in Figure 2 depict the recognition performances in quiet by the listeners with hearing loss at 60 and 80 dB HL, which correspond to the presentation levels of the words in the 0 and 20 dB S/N WIN conditions (the same data for the listeners with normal hearing were >98% correct and are not plotted in Figure 2). At 80 dB HL, the performances by the listeners with hearing loss were essentially the same in quiet (84.7%) and in multitalker babble (86.7%) that indicates the babble provided no masking or

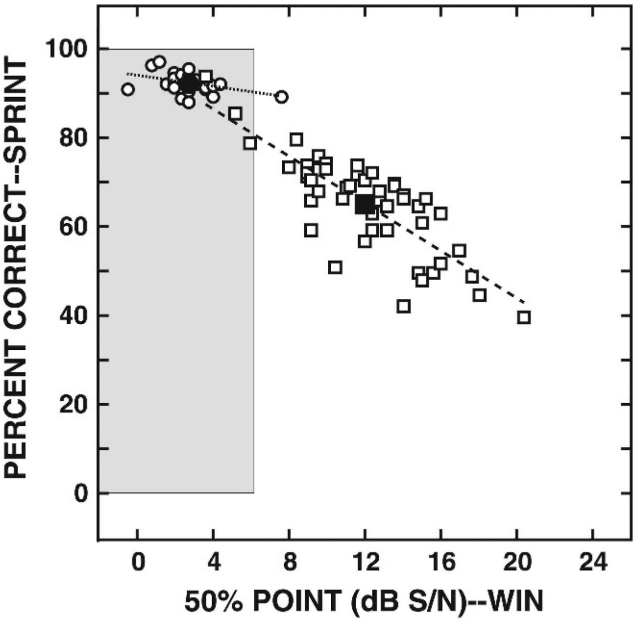

Figure 3. Bivariate plot of the recognition performance on the SPRINT (ordinate) and on the WIN (abscissa) for the 24 listeners with normal hearing (circles) and the 48 listeners with hearing loss (squares). The dashed lines represent the linear functions used to describe the data, and the large filled symbols represent the mean datum points. The shaded region of the graph defines normal performance on the WIN (Wilson, Abrams, et al, 2003).

interference at the 20 dB S/N condition. In contrast, when the presentation level was 60 dB HL, recognition performance decreased in quiet slightly to 75.7%; however, in babble at the same speech presentation level, performance decreased dramatically to 0.2%. Because performance in quiet was so relatively good at 60 dB HL, the lack of performance in babble can not be attributed to the audibility factor but, rather, is accounted for by the detrimental effects caused by the multitalker babble. Likewise, at 60 dB HL, audibility did not account for the poor performance by the listeners with normal hearing whose performance dropped from 98% in quiet to 33.8% in babble.

Finally, from Figure 2, consider the relation between the performances by the listeners with hearing loss on the SPRINT and the WIN shown by the horizontal separation between the large square and the WIN function. On the SPRINT, 65% correct was achieved at 9.0 dB S/N, whereas on the WIN the 65% point was achieved at 13.6 dB S/N (calculated from the polynomial equation), which is a 4.6 dB difference. Based on data obtained in quiet in earlier investigations, this 4.6 dB difference is attributable to speaker differences between the two recordings used by the SPRINT and WIN paradigms. Wilson et al (1990, figure 3) analyzed the psychometric functions generated by four recordings (different speakers) of the NU-6 lists in quiet, including the Auditec recording (used with the SPRINT) and the VA recording (used with the WIN). Although the slopes of the functions were similar, the displacement between the performances on the Auditec and VA recordings of NU-6 in the Wilson et al study differed by about 5 dB, with the VA recording requiring a higher presentation level for equal performance. This is the same relation between the two recorded versions of NU-6 that was observed in the current study.

Figure 3 is a bivariate plot of the performances by the individual participants on the SPRINT (ordinate) and WIN (abscissa). The shaded region of the graph defines normal performance on the WIN (≤6 dB S/N; Wilson, Abrams, et al, 2003). The WIN data are the 50% points derived with the Spearman-Kärber equation that produced means (larger, filled symbols) of 2.7 and 12.0 dB S/N for the listeners with normal hearing and listeners with hearing loss, respectively (SDs = 1.5 and 3.4 dB). With one exception, the listeners with normal hearing are grouped closely above 90% correct on the SPRINT and in the normal range of the WIN. Because the scales on the ordinate and abscissa are different, the apparent inverse relation between performances on the SPRINT and WIN for the listeners with hearing loss in reality indicates that as performances improved on the SPRINT, performances likewise improved on the WIN. Three of the listeners with hearing loss had performances on the WIN that were within the normal range, whereas one of the listeners with normal hearing was outside the normal range. A Pearson product-moment correlation indicated a significant relationship ($r = -0.81$, $p < .01$) between the performances on the SPRINT and WIN. This highly related correlation indicates good concurrent validity of the SPRINT (Franzblau, 1958), which was one of the objectives of this project.

Pearson product-moment correlations indicated significant relationships between performances on the NU-6 materials presented at 80 dB HL and on the SPRINT ($r = 0.71, p < .01$) and on the WIN ($r = -0.75$, $p < .01$). This marked relation for the SPRINT data is illustrated in Figure 4, which is a bivariate plot of the performance on the SPRINT (ordinate) versus performance on NU-6 presented at 80 dB HL in quiet (abscissa) by the listeners with normal hearing (circles) and by the listeners with hearing loss (squares). The large, filled symbols depict the mean data, and the dashed line is the linear regression used to describe the data. The data from the listeners with normal hearing are near maximum for both conditions. In general, for the listeners with hearing loss there is a direct relation between performances on the SPRINT and NU-6 in quiet, but the average difference between the two measures was 19.4%. Although several listeners with hearing loss had 90–100% scores in quiet, most of the SPRINT scores were <80%.

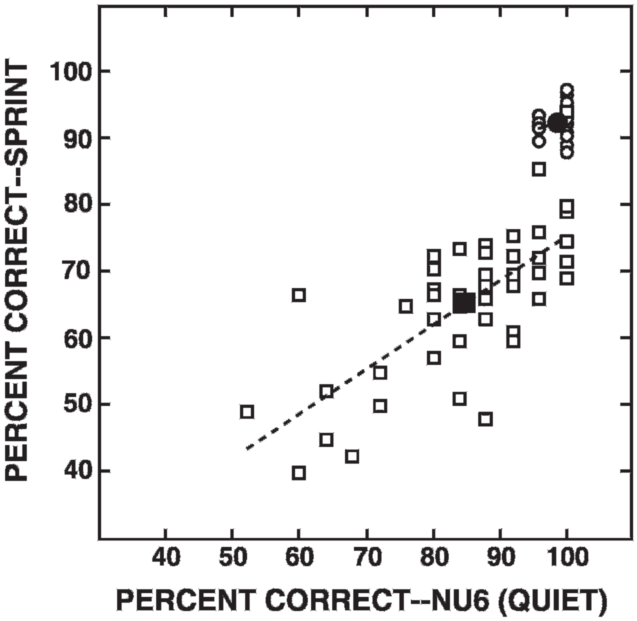

Figure 4. A bivariate plot of the recognition performance on the SPRINT (ordinate) and on the NU-6 in quiet (abscissa) by the 24 listeners with normal hearing (circles) and the 48 listeners with hearing loss (squares). The dashed lines represent the linear functions used to describe the data, and the large filled symbols represent the mean datum points.

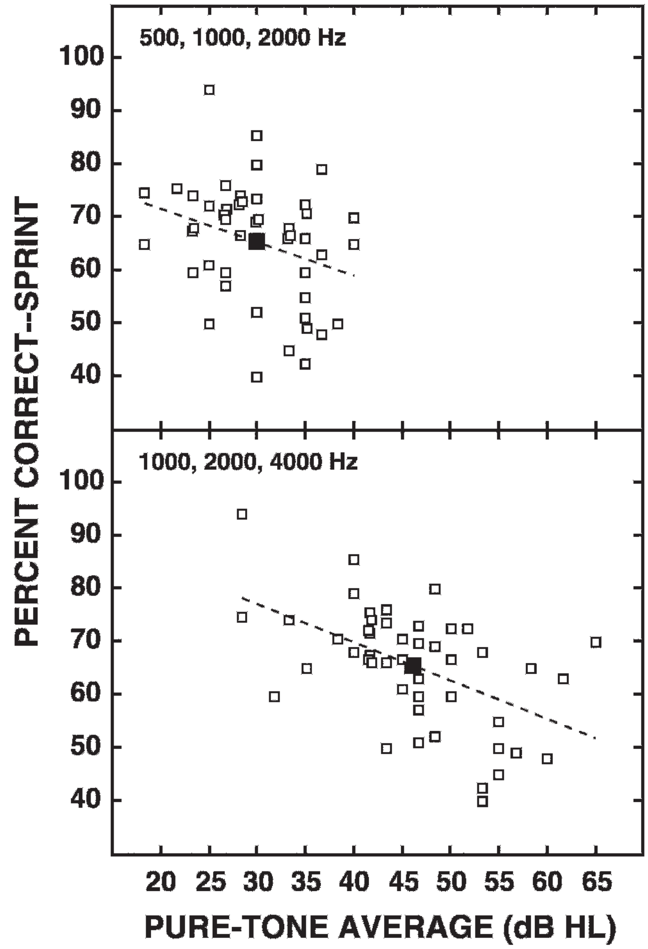

Figure 5. Bivariate plots of performance on the SPRINT (ordinate) by the 48 listeners with hearing loss and the average pure-tones (abscissa) for 500, 1000, and 2000 Hz (top panel) and 1000, 2000, and 4000 Hz (bottom panel). Overlapping datum points were jittered to avoid overlapping datum points. The dashed lines are the linear regressions used to describe the data, and the larger, solid symbols represent the mean data.

The relations between the traditional pure-tone average (PTA at 500, 1000, and 2000 Hz) and performance on the SPRINT and between the high-frequency, pure-tone average (HFPTA at 1000, 2000, and 4000 Hz) and performance on the SPRINT are depicted in Figure 5. Two relations are apparent from the data in the figure. First, as expected from the audiogram in Figure 1, the distribution of datum points and the mean for the HFPTA variable are displaced to the higher hearing levels than are the datum points and mean for the traditional PTA. Second, the slope of the linear regression used to describe the data for the HFPTA variable is steeper than the slope of the function for the traditional PTA. A steeper slope indicates a better relation between the two variables under analysis. The second point is supported by the Pearson product-moment correlations that for the SPRINT were significant, but at different levels, for both pure-tone averages. The SPRINT and the traditional PTA had a fair relation ($r = -0.31$, $p < .05$), whereas the SPRINT and the HFPTA had a moderate relation ($r = -0.52$, $p < .01$). Although not illustrated, the relations between the WIN data and the two pure-tone averages were different from the relations observed with the SPRINT. There was not a significant relation between the WIN and the traditional PTA, whereas there was a moderate relation ($r = 0.56$, $p < .01$) between the WIN and the HFPTA. This was an expected result, as a number of earlier studies have observed this differential relation between word recognition in noise and the two pure-tone averages (e.g., Kryter et al, 1962; Harris, 1965; Wilson and McArdle, 2005, figure 5). Interestingly, the traditional PTA was not significantly related to understanding speech in quiet, whereas the HFPTA and understanding speech in quiet were moderately related ($r = -0.43$, $p < .01$).

Figure 6 contains histograms of the number of the 200 SPRINT words whose percent correct was in each of ten percent-correct intervals. The histograms for the listeners with normal hearing (top panel) and with hearing loss (bottom panel) are shown. As overall performance on the SPRINT by the listeners with normal hearing was 92.5% correct, the majority of words that were correct were in the 90–100% interval. The distribution of correct words with the listeners with hearing loss was quite different with an almost flat distribution across the percent correct intervals from 30–100%. The maximum number of words in any interval was 40 in the 80–90% interval. This unimodal

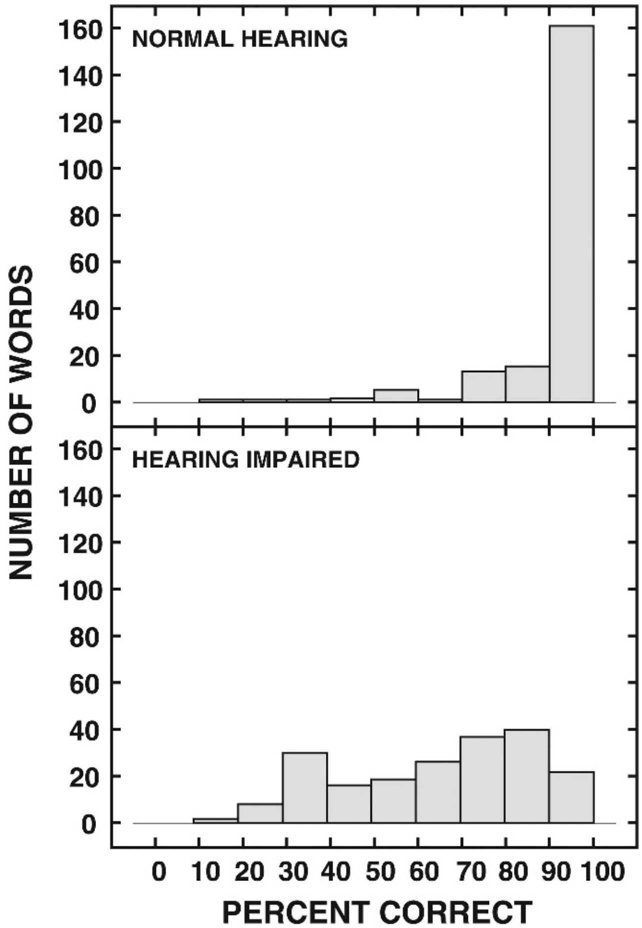

Figure 6. Histograms of the number of SPRINT words identified correctly for the various percent correct categories determined with the 24 listeners with normal hearing (top panel) and the 48 listeners with hearing loss (bottom panel).

distribution of words correct at a given presentation level helps explain the intersubject variability that is associated with word-recognition scores. No doubt, had the data from the listeners with normal hearing been obtained at a less favorable S/N at which recognition performance was 40–60% correct, a similar distribution of words correct would have been observed.

SUMMARY AND CONCLUSIONS

The results from the current study are summarized as follows:

1. Recognition performance on the SPRINT by young listeners with normal hearing (92.5%) was close to the 95–100% performance by listeners with normal hearing targeted by Cord et al (1992).
2. The older listeners with sensorineural hearing loss included in this study performed poorer on the SPRINT (65.3%) than did the younger group of Army personnel with H-3 profiles (81.5%) reported by Cord et al.
3. Graphically and numerically the SPRINT and WIN were highly related, which is indicative of good concurrent validity of the SPRINT.
4. Recognition performances on the NU-6 in quiet and the SPRINT and on the NU-6 in quiet and the WIN were correlated significantly.
5. The HFPTA (1000, 2000, and 4000 Hz) had higher correlations with recognition performance on the SPRINT, WIN, and NU-6 in quiet than did the traditional PTA (500, 1000, and 2000 Hz).

Acknowledgments. 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. The authors acknowledge the contributions made by to the project by Kelly Koder, Rachel McArdle, Sherri Smith, and Brian Walden.

REFERENCES

American National Standards Institute (ANSI). (2004) *Specification for Audiometers* (ANSI S3.6 2004). New York: ANSI.

Cord MT, Walden BE, Atack RM. (1992) Speech Recognition In Noise Test (SPRINT). Walter Reed Army Medical Center.

Department of the Army. (2007) *Standards of Medical Fitness.* Army Regulation 40-501. Washington, DC: Department of the Army. http://www.army.mil/usapa/epubs/pdf/r40_501.pdf.

Department of Veterans Affairs. (1952) *Speech Recognition and Identification Materials.* Disc 4.0. Mountain Home, TN: VA Medical Center.

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

Franzblau AN. (1958) *A Primer of Statistics for Non-statisticians.* New York: Harcourt, Brace and World.

Harris JD. (1965) Pure-tone acuity and the intelligibility of everyday speech. *J Acoust Soc Am* 37:824–830.

Holma T, Laitakari K, Sorri M, Winblad I. (1997). New speech-in-noise test in different types of hearing impairment. *Acta Otolaryngol Suppl* 529:71–73.

Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. (2004) Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am* 116:2395–2405.

Kryter KD, Williams C, Green DM. (1962) Auditory acuity and the perception of speech. *J Acoust Soc Am* 34:1217–1223.

McArdle RA, Wilson RH, Burks CA. (2005) Speech Recognition in Multitalker Babble Using Digits, Words, and Sentences. *J Am Acad Audiol* 16:726–739.

Nilsson M, Soli S, Sullivan J. (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.

Niquette P, Arcaroli J, Revit L, Parkinson A, Staller S, Skinner M, Killion M. (2003) Development of the BKB-SIN Test. Paper presented at the Annual Meeting of the American Auditory Society, Scottsdale. http://www.amauditorysoc.org/pages/abstracts/2003/2003-pp-1.html.

Sweetow RW, Sabes JH. (2006) The need for the development of an adaptive Listening And Communication Enhancement (LACE^{TM}) Program. *J Am Acad Audiol* 17:538–558.

Thornton AR, Raffin MJM. (1978) Speech-discrimination scores modeled as a binomial variable. *J Speech Hear Res* 21:507–518.

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

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, Bell TS, Koslowski JA. (2003) Learning effects associated with repeated word-recognition measures using sentence materials. *J Rehabil Res Dev* 40:329–336. http://www.vard.org/jour/03/40/4/pdf/Wilson-A.pdf.

Wilson RH, Burks CA. (2005) The Use of 35 Words to Evaluate Hearing Loss in Terms of Signal-to-Babble Ratio: a Clinic Protocol. *J Rehabil Res Dev* 42:839–852. http://www.rehab.research.va.gov/jour/05/42/6/pdf/wilson.pdf.

Wilson RH, Coley KE, Haenel JH, Browning KM. (1976) Northwestern University Auditory Test No. 6: Normative and Comparative Intelligibility Functions. *J Am Audiol Soc* 1:221–228.

Wilson RH, McArdle R. (2005) Speech signals used to evaluate auditory function. *J Rehabil Res Dev* 42(Suppl. 2):79–94.

Wilson RH, McArdle R. (2007) Intra- and inter-session test, retest reliability of the Words-in-Noise (WIN) Test. *J Am Acad Audiol* 18:819–831.

Wilson RH, McArdle RA, Smith SL. (2007) An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *J Speech Lang Hear Res* 50:844–856.

Wilson RH, Zizz CA, Shanks JE, Causey JD. (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.