Journal of Rehabilitation Research and Development
Vol. 40, No. 4, July/August 2003
Pages 321–328

A word-recognition task in multitalker babble using a descending presentation mode from 24 dB to 0 dB signal to babble

Richard H. Wilson, PhD; Harvey B. Abrams, PhD; Amanda L. Pillion, MEd
James H. Quillen Department of Veterans Affairs (VA) Medical Center, Mountain Home, TN; Departments of Surgery and Communicative Disorders, East Tennessee State University, Johnson City, TN; VA Medical Center, Bay Pines, FL; Department of Communication Sciences and Disorders, University of South Florida, Tampa, FL

Abstract—A speech-in-multitalker-babble test instrument was developed for use in a Department of Veterans Affairs (VA) multicenter study examining the effects of hearing loss on self-perceived quality of life. Word recognition in quiet and in multitalker babble was measured on 24 listeners with normal hearing and 24 listeners with sensorineural hearing loss. The protocol involved the presentation of 10 monosyllabic words (each in a unique babble segment) at each of seven signal-to-babble (S/B) ratios from 24 dB to 0 dB, with the babble fixed at 60 dB HL (hearing loss). Word recognition in quiet at 60 dB and 80 dB HL for both groups was >90% correct. Two trials on the task were conducted. In babble, the 50% correct points were at 4.1 dB and 9.4 dB S/B for the listeners with normal hearing and hearing loss, respectively, with the 90th percentile for the listeners with normal hearing at 6 dB S/B. Twenty-two of the twenty-four listeners with hearing loss had 50% correct points outside of the 90th percentile for listeners with normal hearing. Test-retest reliability was excellent.

Key words: auditory perception, hearing loss, speech perception, word recognition in multitalker babble.

INTRODUCTION

Patients with hearing loss often complain that they can hear but cannot understand speech, especially in the presence of background noise [1–11] Carhart and Tillman and later Plomp and Duquesnoy emphasized that listeners with peripheral sensorineural hearing loss are most handicapped when listening in the presence of background noise [2,12]. In routine clinical practice, the ability of patients to understand speech in a noisy environment is not typically assessed [13,14]. This paper describes a prototype speech-in-noise paradigm in which word-recognition ability is measured in a background of multitalker babble. With this speech-in-babble paradigm, the hearing loss for speech in terms of decibels signal-to-babble (S/B) ratio can be established.

The test instrument described was designed specifically as a component of the auditory evaluation protocol for patients with hearing loss included in a Department of Veterans Affairs (VA) multicenter project [15]. This project measured the effects of audiological intervention on communication function and health-related quality of life as measured by selected generic and disease-specific outcome measures.

Abbreviations: ANOVA = analysis of variance, HL = hearing loss, NU 6 = Northwestern University Auditory Test No. 6, S/B = signal to babble.

This material was based on work supported by a multicenter merit review (C00-1892RA) from the Rehabilitation Research and Development Service to the second author. The first author is a senior research career scientist sponsored by the Rehabilitation Research and Development Service, Department of Veterans Affairs.

Address all correspondence and requests for reprints to Richard H. Wilson, PhD; VA Medical Center, Audiology (126), Mountain Home, TN 37684; 423-926-1171, ext. 7553; fax: 423-979-3403; richard.wilson2@med.va.gov.

As a means of evaluating speech understanding in background noise for individuals with sensorineural hearing loss, a first approximation of a test instrument was developed in a series of earlier studies [16–18]. The methods and procedures associated with the development of the test materials used in the current study are described in detail elsewhere [19]. Briefly, however, to meet the requirements of the multicenter study, the test instrument had the following characteristics:

1. The Northwestern University Auditory Test No. 6 (NU 6) monosyllabic materials recorded on a disk by a VA female speaker were selected, thereby providing a measure of speech understanding in both quiet and background noise using the same speaker and speaking the same words [20].
2. Multitalker babble was selected as the competing background noise because multitalker babble is the most common environmental noise encountered by listeners in everyday life and because babble is more detrimental to speech perception than are other types of competition [21]. The multitalker babble, which was recorded by Causey in 1988,* consists of three female and three male speakers talking simultaneously about various topics with none of the conversations intelligible [22].
3. Because background noises are maintained at fairly constant levels from one setting to the next, the level of the multitalker babble was fixed and the level of the speech signal varied.
4. The instrument evolved as one in which 10 words were presented at each of seven S/B ratios (24 dB to 0 dB) in a quasi-randomized design. The use of multiple S/B ratios provides the shape of the recognition function. For listeners with normal hearing, the target performance outcomes in percent correct at 0, 4, and 8 dB S/B were 30, 50, and 70 percent, respectively, with performance >90 percent correct at levels above 8 dB S/B. The extension of the S/B ratios in the test paradigm to 12, 16, 20, and 24 dB S/B provided an ample range over which the performances of listeners with hearing loss could be measured.
5. The instrument design was amenable to quantification by the Spearman-Kärber equation that is a simple metric yielding an estimate of the 50 percent correct point in terms of the decibel S/B ratio [23].

*Causey GD. Personal communication, 1988.

In the development of the words in multitalker babble, pilot data from a larger project revealed that performance on a given word was different in different segments of babble [24]. To reduce performance variability on the words in babble, we fixed the temporal relationship between the target word and the multitalker babble segment for that word. Thus, each word was time-locked in a unique 6 s segment of babble. The words were digitally adjusted in level and mixed with the babble segments. To produce the test lists, we shuffled and concatenated the babble segments containing the words. To make the boundaries between babble segments acoustically and perceptually transparent, we edited the onsets and offsets of the waveforms at negative going zero crossings. Through a series of pilot studies, psychometric functions were developed for each of the 150 words in Lists 2, 3, and 4 of NU 6. Based on data from both listeners with normal hearing and listeners with peripheral hearing loss, the pool of words was reduced to 70 words that were sorted into seven S/B ratios based on the recognition performance of the listeners (i.e., 10 words at each S/B ratio). Two versions of the test materials were developed, one in which the 70 words were presented in a randomized S/B ratio paradigm and one in which the words were presented in a descending S/B ratio paradigm [24]. The latter design was selected for use in the multicenter project. In the descending-level paradigm, the 10 words at 24 dB S/B were presented first, followed by 10 different words at 20 dB S/B, etc. In this manner, the listening task progressed from easy to difficult, thereby maximizing the effects, if any, of learning to listen in a background noise.

This study established normative data for the target words used in the multitalker babble paradigm that was incorporated into the multicenter study. It also examined the test-retest reliability of the test instrument. Listeners with sensorineural hearing loss were included to ensure that the protocol would provide appropriate data for listeners with hearing loss.

METHODS

Materials

The two randomizations of the word lists in multitalker babble were recorded on an audio compact disc (Matshita, Model UJDA710) along with Lists 1 and 2 of the NU 6 in quiet. The same female speaker recorded all materials.

Subjects

A group of 24 listeners (mean = 21.1 years) with normal hearing (≤20 dB HL at 250 Hz to 8,000 Hz) and 24 listeners (mean = 58.5 years) with mild-to-moderate sensorineural hearing loss participated. The subjects with hearing loss had (1) word-recognition scores in quiet of ≥76 percent correct at 50 dB HL, (2) pure-tone thresholds that were symmetrical (±10 dB), and (3) ipsilateral acoustic reflexes at 500, 1,000, and 2,000 Hz. The mean audiogram for the listeners with hearing loss can be seen in **Figure 1**. The subjects were remunerated for their participation in the 1-hour session.

Procedures

Each subject was presented NU 6 Lists 1 and 2 in quiet. The odd numbered subjects listened to List 1 at 80 dB HL followed by List 2 at 60 dB HL. The even numbered subjects listened to List 1 at 60 dB HL followed by List 2 at 80 dB HL. The 60 dB and 80 dB HL levels were selected for the quiet presentations, as they were the same levels of the words in multitalker babble at 0 dB and 20 dB S/B, respectively. Following the two quiet lists, each listener was presented two trials on the words in multitalker babble paradigm. The odd numbered subjects were presented the first randomization followed by the second randomization with the even number subjects listening to the reverse sequence. The binaurally presented materials were reproduced on a compact disc player (Sony, Model CDP-497), routed through an audiometer (Grason-Stadler, Model 10) to a pair of TDH-50P earphones encased in Telephonics P/N 510C017-1 cushions. Binaural presentations were used because (1) monaural and binaural data from an earlier study indicated only about a 1 dB difference between performances on the two conditions [19] and because (2) the larger merit review project for which the materials were designed required a binaural measure of word recognition in background noise. For all conditions, the level of the multitalker babble was fixed at 60 dB HL and the level of the speech varied from 60 dB to 84 dB HL. All testing was conducted in a double-wall sound booth. The verbal responses of each listener were recorded into a spreadsheet, which permitted analysis of the individual responses.

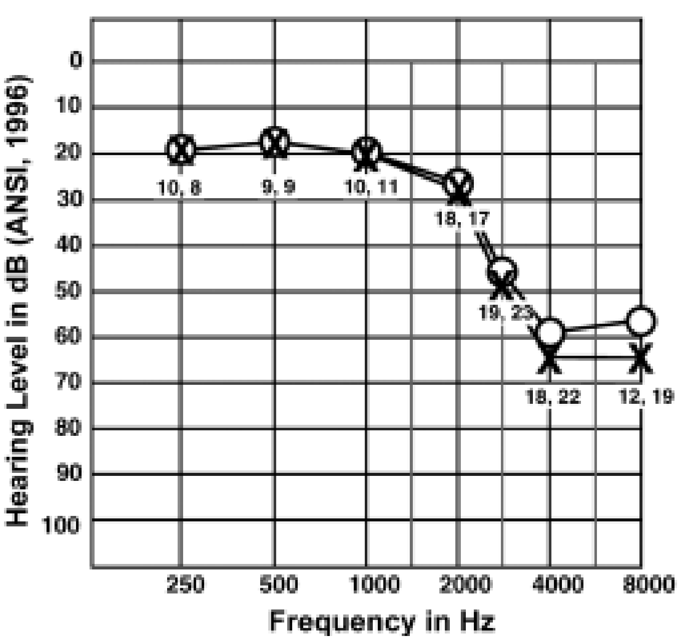

Figure 1.
Pure-tone audiograms for left ears (Xs) and right ears (Os) of 24 listeners with hearing loss included in study. Numbers below each set of symbols are threshold SDs for left and right ears, respectively.

RESULTS AND DISCUSSION

The mean percent correct data (and standard deviations [SDs]) for Lists 1 and 2 of NU 6 presented in quiet are listed in **Table 1**. Although the recognition performances were near maximum, a two-within (list by level), one-between (group) analysis of variance (ANOVA) indicated that recognition performance by the listeners with hearing loss was significantly poorer (<3 tokens) than the recognition performance by the listeners with normal hearing [$F(1,22) = 27.363$, $p < 0.0001$] [25]. The differences between performances on List 1 and List 2 were smaller (<2 tokens) and were not significant ($p > 0.01$).

Table 2 lists the percent correct recognition obtained on the 10 words in quiet from Lists 1 and 2 of NU 6 that

Table 1.
Means and SDs for percent correct word recognition on Lists 1 and 2 of NU 6 in quiet at 60 dB and 80 dB HL.

dB HL	List 1		List 2	
	Mean	SD	Mean	SD
	Listeners with Normal Hearing			
60	99.0	3.2	98.0	5.5
80	97.8	5.5	98.3	5.3
	Listeners with Hearing Loss			
60	94.0	8.9	91.3	10.0
80	95.7	8.6	92.7	9.6

Table 2.
Percent correct recognition for 10 most often incorrect words on Lists 1 and 2 of NU 6 for listeners with normal hearing and for listeners with hearing loss. Percentages represent average obtained for each group in quiet at 60 dB and 80 dB HL.

List 1	Normal Hearing	Hearing Loss	List 2	Normal Hearing	Hearing Loss
Death	88	63	Calm	71	54
Mode	83	79	Dab	100	75
Pool	92	75	Nice	83	92
Met	92	79	Turn	100	75
Laud	92	88	Mill	100	79
Dime	100	83	Pick	92	88
Sell	92	92	South	100	79
Choice	96	88	Bite	100	83
Keen	100	88	Merge	100	83
Sub	100	88	Lore	96	88

were most often incorrect. The values are the mean percents correct obtained at 60 dB and 80 dB HL. The words are ranked according to the overall percent correct recognition obtained by the two groups of listeners. Although performances were generally above 90 percent correct for listeners in both groups, the point of the listings in **Table 2** is to highlight the percent correct variability both among the words and between the two groups. For example, most audiologists would probably guess that the word "laud" in List 1 would be the word most often incorrect. In List 1, however, in quiet four other words (death, mode, pool, and met) were more often incorrect than "laud." For both lists, the word "calm" was the most often missed word. No doubt, the rankings of word difficulty in **Table 2** will vary depending on the ranges of percent correct recognition included in the sample. The listeners with normal hearing performed better on 58 percent of the 100 words with equal performance by the two groups on 37 percent of the words. The listeners with hearing loss performed minimally better than did the listeners with normal hearing on 5 percent of the words.

The mean percents correct for the words in multitalker babble from the two groups on the two trials are listed in **Table 3**. The SDs for both the subjects and the words are listed. Also included for the four conditions are the overall percent correct and the 50 percent correct points calculated with the Spearman-Kärber equation and expressed in terms of both the decibel S/B ratio and the decibel hearing level of the words [26]. A one-within (trial), one-between (group) ANOVA on the 50 percent points revealed that the only significant difference was between groups [$F(1,46) = 52.556$, $p < 0.0001$]. As expected, the intersubject SDs indicate greater variability among the listeners with hearing loss than among the listeners with normal hearing. For both groups, the largest variability was on the dynamic segment of the psychometric function. Interestingly, the interword variability on the dynamic segment of the function was smaller for the listeners with hearing loss than for the listeners with normal hearing.

The excellent agreement between the mean 50 percent correct points on the two trials is illustrated in **Figure 2** with the data from the listeners with normal hearing (□) and the listeners with hearing loss (○). **Figure 2** is a bivariate plot of the test (abscissa) and retest (ordinate) points calculated with the Spearman-Kärber equation. The solid symbols depict the mean data. The numbers on the graph indicate the number of data points above, on, or below the diagonal line that represents equal 50 percent correct points on the two trials. Nine of the listeners with hearing loss had 50 percent points that were at higher S/B ratios on Trial 2 than on Trial 1. Five of the listeners with hearing loss had equal performances on the two trials. The shaded region in the lower left corner of the figure represents the 90th percentiles (6 dB S/B) for Trial 1 and Trial 2 from the listeners with normal hearing. As illustrated, the levels of the 50 percent points for the listeners with normal hearing are clustered closely, whereas the levels for the 50 percent points for the listeners with hearing loss are spread widely, with most outside of the 90th percentile area. In fact, performances by 22 of the 24 listeners with hearing loss were outside of the 90th percentile region. Finally, in the original design of the test instrument, the target percents correct at 0, 4, and 8 dB S/B for the listeners with normal hearing were 30, 50, and 70 percent. The data at these S/B ratios in **Table 2** for the listeners with normal hearing are very close to the target values. Because no significant difference was found between the data from Trial 1 and Trial 2, the remaining data presented and discussed are for Trial 1.

The psychometric functions for the listeners with normal hearing (□) and with hearing loss (○) on Trial 1 are presented in **Figure 3**. The S/B ratios at which the 50 percent correct recognition performances occurred on the mean functions in **Figure 3** were 2.9 dB and 8.4 dB for the listeners with normal hearing and the listeners with hearing loss, respectively. Thus, the listeners with hearing

Table 3.
Mean percent correct word recognition obtained at respective S/B ratios (dB S/B) and at presentation levels (dB HL) of words. SDs are listed for subjects and for words. Data listed for 50 percent correct points were derived from Spearman-Kärber equation. Babble was constant at 60 dB HL.

dB S/B	dB HL	Trial 1			Trial 2		
		Mean	$SD_{subject}$	SD_{word}	Mean	$SD_{subject}$	SD_{word}
		Listeners with Normal Hearing					
24	84.0	97.5	4.4	7.9	97.5	4.4	7.9
20	80.0	99.6	2.0	1.3	99.2	2.8	1.8
16	76.0	98.3	3.8	2.2	99.2	2.8	1.8
12	72.0	98.8	3.4	2.0	98.8	3.4	2.0
8	68.0	72.9	11.2	24.3	74.6	13.5	23.0
4	64.0	56.3	17.4	25.6	58.3	19.3	27.1
0	60.0	23.3	14.3	16.3	23.8	16.1	18.1
Overall % Correct	78.1	5.0	—	78.8	4.8	—	—
50% Point (dB S/B)	4.1	1.4	—	4.0	1.3	—	—
50% Point (dB HL)	64.1	—	—	64.0	—	—	—
		Listeners with Hearing Loss					
24	84.0	90.4	16.8	8.8	94.2	11.4	3.5
20	80.0	95.4	11.0	5.4	92.1	16.1	7.5
16	76.0	84.6	17.4	13.0	87.5	16.2	8.8
12	72.0	82.5	17.0	10.7	81.7	16.9	12.1
8	68.0	47.5	25.7	16.8	42.5	26.9	13.4
4	64.0	13.8	15.8	8.6	12.9	18.8	8.7
0	60.0	1.7	4.8	2.9	0.8	2.8	2.6
Overall % Correct	59.4	12.0	—	60.2	15.9	—	—
50% Point (dB S/B)	9.4	3.4	—	9.1	4.5	—	—
50% Point (dB HL)	69.4	—	—	69.1	—	—	—

loss required a 5.5 dB enhanced S/B ratio to obtain recognition performance that was equivalent to the recognition performance of the listeners with normal hearing. A 5 dB HL for pure tones creates little, if any, deficit in auditory function. In contrast to a 5 dB sensitivity hearing loss, a 5 dB HL in terms of S/B ratio is devastating and difficult to overcome. These disparate measures of speech recognition performance obtained in quiet and in background noise are found throughout the literature [1–11]. The slopes of the functions at the 50 percent correct points were 6.5%/dB and 4.5%/dB, respectively. The more gradual slope of the function for listeners with hearing loss reflects the larger intersubject variability associated with this group. The levels of the 50 percent points derived from the polynomials agree with the levels of the 50 percent points calculated with the Spearman-Kärber equation on the Trial 1 data (4.1 dB and 9.4 dB in **Table 3**). For comparison, the filled symbols at 60 dB and 80 dB HL depict recognition performance on the NU 6 materials in the quiet condition for the two groups. The relationship between the datum points at 60 dB HL indicates that the poorer performance by both groups of listeners in the multitalker babble is the result of the "masking" (or distortion in the Plomp model) created by the babble [27], not the result of poor word recognition in quiet. Finally, from the Trial 1 data in **Table 3**, the 18.7 percent (or 5.3 dB) difference between the performances in multitalker babble by the two groups of listeners (78.1% to 59.4%) illustrates how much more of an adverse affect background noise has on listeners with hearing loss than on young adults with normal hearing. The subjects with hearing loss included in this study were ideal subjects; i.e., they had good word-recognition abilities in quiet at 50 dB HL. Based on previous and current data, an inverse relation is

Journal of Rehabilitation Research and Development Vol. 40, No. 4, 2003

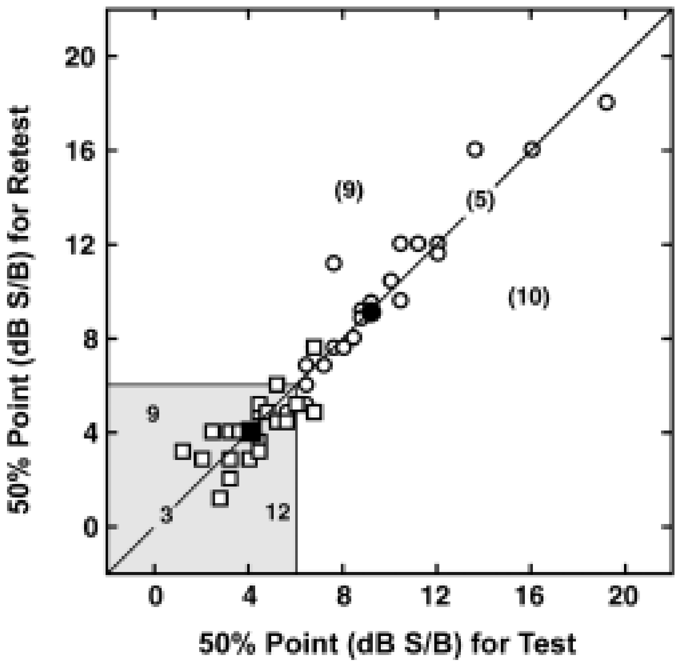

Figure 2.
Decibel S/B levels of 50% correct points on functions for 24 listeners with normal hearing (□) and 24 listeners with hearing loss (○). Large filled symbols represent mean data. Shaded area in lower left defines 90th percentile for group with normal hearing. Numbers for two groups above, on, and below diagonal line indicate number of listeners whose 50% points were higher on Trial 2 than on Trial 1, equal on two trials, and higher on Trial 1 than on Trial 2, respectively.

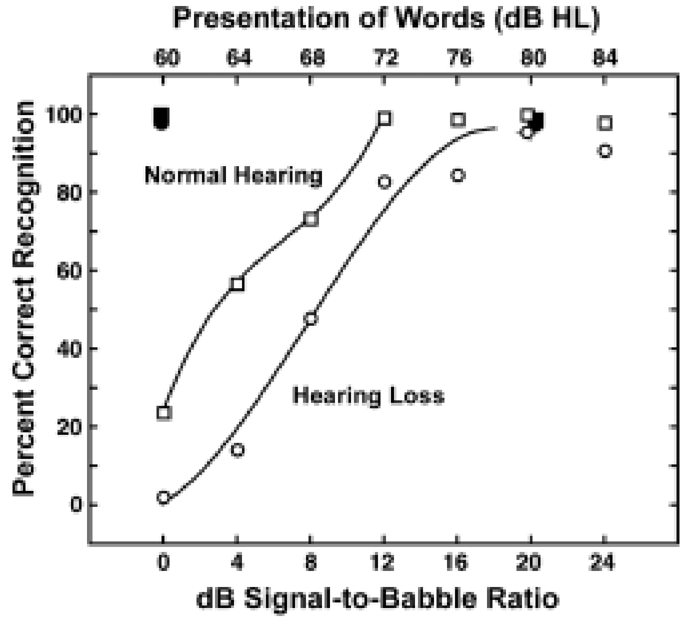

Figure 3.
Mean percent correct word recognition on Trial 1 for 24 listeners with normal hearing (□) and for 24 listeners with hearing loss (○). Lines connecting datum points are best-fit third-degree polynomials used to fit dynamic portions of functions for listeners with normal hearing [$Y = 23.300 + (12.4417x) + (-1.3156x^2) + (0.0669x^3)$; $r^2 = 1.0$] and listeners with hearing loss [$Y = -1.3786 + (4.7440x) + (0.2710x^2) + (-0.0130x^3)$; $r^2 = 0.98$]. Percent correct recognition on NU 6 materials in quiet at 60 dB and 80 dB HL is depicted as filled symbols.

expected both between degree of peripheral hearing loss and word-recognition performance in multitalker babble and between age and word-recognition performance in multitalker babble. Poorer word-recognition performance in multitalker babble is anticipated from listeners with peripheral hearing losses greater than those included in this study and from listeners who are older than those included in this study.

CONCLUSIONS

The data reported in this study indicate that the word-recognition task in the multitalker babble paradigm developed for the multicenter study is a technique that can efficiently quantify the ability of individual listeners with either normal hearing or hearing loss to understand speech in background noise. The measure defines an hearing loss for speech in terms of the S/B ratio. The listeners with normal hearing and the listeners with peripheral hearing loss had similar word-recognition abilities in quiet at 60 dB and 80 dB HL. For equal recognition performance in multitalker babble, however, listeners with mild-to-moderate hearing loss required on average a 5.5 dB more favorable S/B ratio than did listeners with normal hearing. The 90th percentile for the listeners with normal hearing was 6 dB S/B. Only 2 of the 24 listeners with hearing loss were within this 90th percentile. No difference was found between the multitalker babble data from the two trials indicating excellent test-retest reliability. The data obtained with the words in multitalker babble paradigm clinically demonstrate the difficulty that listeners with hearing loss have when listening in background noise.

REFERENCES

1. Hirsh IJ. Binaural hearing aids: A review of some experiments. J Speech Hear Disord 1950;15:114–23.

2. Carhart R, Tillman TW. Interaction of competing speech signals with hearing loss. Arch Otolaryngol 1970;91:273–79.
3. Keith RW, Talis HP. The use of speech in noise in diagnostic audiometry. J Auditory Res 1970;10:201–4.
4. Olsen WO, Noffsinger D, Kurdziel S. Speech discrimination in quiet and in white noise by patients with peripheral and central lesions. Acta Otolaryngol 1975;80:375–82.
5. Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am 1984;76:87–96.
6. Gordon-Salant S. Age-related differences in speech recognition performance as a function of test format and paradigm. Ear Hear 1987;8:277–82.
7. Beattie RC. Word recognition functions for the CID W-22 test in multitalker noise for normally hearing and hearing-impaired subjects. J Speech Hear Disord 1989;54:20–32.
8. Souza PE, Turner CW. Masking of speech in young and elderly listeners with hearing loss. J Speech Hear Res 1994;37:655–61.
9. Divenyi PL, Haupt KM. Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. I. Age and lateral asymmetry effects. Ear Hear 1997;18:42–61.
10. Divenyi PL, Haupt KM. Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. II. Correlation analysis. Ear Hear 1997;18:100–13.
11. Divenyi PL, Haupt KM. Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation. Ear Hear 1997;18:189–201.
12. Plomp R, Duquesnoy AJ. A model for the speech-reception threshold in noise without and with a hearing aid. Scand Audiol 1982;15:95–111.
13. Martin FN, Champlin CA, Chambers JA. Seventh survey of audiometric practices in the United States. J Am Acad Audiol 1998;9:95–104.
14. Mueller HG. Speech audiometry and hearing aid fittings: Going steady or casual acquaintances. Hear J 2001;54:19–29.
15. Abrams HB, Doyle PJ. Functioning, disability, and quality of life in the adult with hearing impairment. Rehabilitation Research and Development Service Merit Review, Department of Veterans Affairs, Washington (DC); 2000.
16. Wilson RH, Sumrall R, Oyler AL. Influence of stimulus length uncertainty on dichotic digits. American Academy of Audiology Convention; 1996 Apr; Salt Lake City (UT).
17. Norwood-Chapman L, Wilson RH, Thelin JW. Clinical evaluation of word recognition in multitalker babble under four listening conditions, American Academy of Audiology Convention; 1997 Apr; Fort Lauderdale (FL).
18. Wilson RH, Strouse A. Northwestern University Auditory Test No. 6 in multitalker babble: A preliminary report. J Rehabil Res Dev 2002;39:105–13. Available from: URL: http://www.vard.org/jour/02/39/1/wilson.htm.
19. Wilson RH. Development of a speech in multitalker babble paradigm to assess word-recognition performance in S/B ratios from 24 dB to 0 dB. J Am Acad Audiol. In press 2003.
20. Department of Veterans Affairs Speech Recognition and Identification Materials, CD 2.0 (1998), Mountain Home, TN: VA Medical Center. Available from: URL: http://www.va.gov/621quillen/clinics/asp/Products/SPEECHCD/sraim-index.htm.
21. Findlay RC. Auditory dysfunction accompanying noise-induced hearing loss. J Speech Hear Disord 1976;41:374–80.
22. Sperry JL, Wiley TL, Chial MR. Word recognition performance in various background competitors. J Am Acad Audiol 1997;8:71–80.
23. Finney DJ. Statistical method in biological assay. London: C. Griffen; 1952.
24. Wilson RH. Evaluation of the hearing loss for speech in multitalker babble. Rehabilitation Research and Development Service Merit Review, Department of Veterans Affairs, Washington (DC); 2002.
25. Abacus Concepts. SuperANOVA, Berkeley CA: Abacus Concepts, Inc., 1989.
26. American National Standards Institute (ANSI). Specification for audiometers (ANSI S3.6-1996). New York: ANSI; 1996.
27. Plomp R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. J Acoust Soc Am 1978;63:533–49.

Submitted for publication October 11, 2002. Accepted in revised form January 8, 2003.