Main paper

Normal and hearing-impaired word recognition scores for monosyllabic words in quiet and noise

Randall C. Beattie, Tiffany Barr and Christina Roup

Department of Communicative Disorders, California State University, Long Beach, Long Beach, CA 90840, USA

(Received 19 December 1995, accepted 11 July 1996)

Abstract
The effects of noise on word recognition scores were assessed with normal-hearing and hearing-impaired subjects. Fifty-one normal-hearing subjects were tested at 50 dB HL using signal-to-noise ratios (S/Ns) of 5, 10, and 15 dB. Thirty subjects with mild-to-moderate sensorineural hearing losses were tested in quiet and in noise at S/Ns of 10 dB and 15 dB. Monosyllabic words in a Multitalker Noise were selected for testing. Mean scores for the normal-hearing subjects were 45% at the 5 dB S/N, 74% at the 10 dB S/N, and 87% at the 15 dB S/N. For the hearing-impaired subjects, scores were 85% in quiet, 60% at the 15 dB S/N, and 40% at the 10 dB S/N. These results suggest that background noise which is mildly disruptive for normal hearing subjects can be highly disruptive to hearing-impaired subjects. Moreover, these findings indicate that subjects with mild-to-moderate sensorineural hearing loss require a more favourable S/N than normal listeners to achieve comparable word recognition scores. Test-retest differences for word recognition scores revealed variability that agreed closely with predictions based on the binomial distribution for both groups of subjects. Speech-in-noise abilities must be measured directly because regression equations revealed that speech-in-noise scores cannot be predicted accurately from either puretone thresholds or speech-in-quiet scores. Word recognition functions are presented from several hearing-impaired subjects and demonstrate the value of testing in noise.

Key words: speech in noise; word recognition; hearing impaired

Audiological tests using speech stimuli are essential when evaluating patients with hearing loss because speech is the most important auditory stimulus in everyday situations. Several authors recommend measuring speech recognition in the presence of a competing noise (Jerger and Hayes, 1976; Plomp and Mimpfen, 1979; Dirks et al., 1982; Plomp, 1986; Beattie, 1989; Moore et al., 1992a,b; Gatehouse, 1994). Speech-in-noise measurements have been obtained for a variety of reasons. First, speech recognition in noise may be used to ascertain which ear is most suitable for amplification (Hagerman, 1984; Penrod, 1994). For example, if both ears exhibit similar audiograms, the clinician may choose to aid the ear that exhibits the highest speech recognition score (Berger et al., 1980; Beattie and Warren, 1982).

Second, performance on speech-in-noise tests may be used to award compensation (Hagerman, 1984). That is, insurance agencies may employ speech-in-noise tests as a more valid measure of communication handicap than puretone or speech-in-quiet tests. Speech-in-noise tests also have been used to compare hearing aids (Jerger and Hayes, 1976; Schwartz et al., 1979; Frank and Craig, 1984; Moore et al., 1992a,b; Flexer, 1993). The hearing aid or hearing aid arrangement that provides the best speech recognition score may be recommended. Third, speech-in-noise tests are used to identify the probable site of auditory pathology (Jerger and Jerger, 1981; Hall, 1983; Stecker, 1992; Mueller and Bright, 1994). For example, Morales-Garcia and Poole (1972) present data in support of speech-in-noise testing for
differentiating lesions of the central auditory pathways. Fourth, some authors recommend that word recognition scores be obtained before and after surgery because both improvement and degradation have been observed (Penrod, 1994). Finally, word recognition tests may provide insight when evaluating social adequacy and communicative effectiveness (Gatehouse, 1990, 1994; Schow and Gatehouse, 1990; Flexer, 1993; Penrod, 1994).

One reason for adding background noise to test stimuli is to make the test more representative of real-life listening (Jerger and Hayes, 1976; Kalikow et al., 1977; Hagerman, 1984; Plomp, 1986; Schow and Gatehouse, 1990). That is, speech-in-noise tests may be used to identify and demonstrate communicative difficulties. Although S/Ns in homes and schools typically vary from about 5 dB to 20 dB (Flexer, 1993; Nabelek and Nabelek, 1994), S/Ns varying from about 5 dB to 5 dB have been measured in public and cocktail party environments (Pearsons et al., 1976; Plomp, 1977; Crandell and Smaldino, 1995). Evaluating speech recognition in quiet may not provide a realistic index of communicative difficulty in everyday situations because they are often characterized by competing noise (Jerger and Hayes, 1976; Dirks et al., 1982; Plomp, 1986; Gatehouse and Haggard, 1987). Moreover, measuring speech recognition in quiet does not predict performance accurately in background noise. Because conversations often occur in the presence of competing signals, and word recognition performance in noise cannot be predicted from performance in quiet, several investigators have suggested that the audiologic battery should include speech-in-noise measurements (Cooper and Cutts, 1971; Beattie, 1989). Hearing-impaired listeners typically exhibit poorer speech recognition in noise than normal-hearing subjects (Carhart and Tillman, 1970; Dirks et al., 1982; Beattie, 1989). Performance on speech-in-noise tests may enable the clinician to provide hearing-impaired individuals or family members with more realistic expectations of unaided and/or aided auditory performance in everyday situations (Jerger and Hayes, 1976; Kalikow et al., 1977). Gatehouse and Haggard (1987, pg. 144) state: 'In the event of only one measure being affordable, speech-in-noise may be the preferable one, as it better indicates the worst-case of disability when communication is not well-managed, and better indicates the rather limited benefit a hearing aid can provide for some individuals.

A second reason for adding noise to speech is to increase the difficulty of the test in an effort to identify differences among hearing losses or hearing aids (Dillon, 1983; Plomp, 1986; Moore et al., 1992a,b). Speech stimuli such as sentences, spondees, and even monosyllables often are too easy to separate normal-hearing from mild hearing losses (Carhart, 1965; Speaks et al., 1967). That is, because scores often approach 100% (ceiling effect) for subjects having mild hearing losses, these tests are not sensitive to the communicative difficulties imposed by hearing loss. Moreover, testing in quiet may not reveal differences between hearing aids (Schwartz et al., 1979).

Loven and Hawkins (1983) state that the addition of noise to speech may change the structure of the test so that whatever ability is measured in quiet may not be the same ability measured when the speech is mixed with noise. Consistent with this statement is the observation that speech recognition errors are differentially affected by the addition of background noise (Dubno and Levitt, 1981). Moreover, masking of a particular stimulus is dependent on the intensity, spectrum, presence and duration of acoustic windows, and the meaningfulness of the competing noise (Kalikow et al., 1977; Dillon, 1983; Duquesnoy, 1983). Kalikow et al. (1977) comment that a speech babble can produce more masking than a speech noise because the babble produces false speech cues and increases the load on attentional and memory processes.

The first experiment was conducted to establish normal performance on a speech-in-noise test so that these data could serve as a reference for patients with mild-to-moderate sensorineural hearing loss. The second experiment was designed to ascertain how subjects with mild-to-moderate sensorineural hearing loss perform in comparison with normal listeners and to gather preliminary data on the contribution of speech-in-noise testing to the management of the hearing-impaired listener. Monosyllabic words in a Multitalker Noise were selected for testing. Preliminary data collected in this laboratory suggested that S/Ns varying from about 5 to 15 dB would be appropriate (Beattie, 1989). That is, an S/N of 5 dB should be sufficiently taxing to subjects with normal-hearing and very mild hearing losses (avoid ceiling effect) and S/Ns of 10-15 dB.
Word recognition in quiet and noise

should not be too difficult for most subjects with mild-to-moderate hearing losses (avoid floor effect). The selected range of S/Ns also is typical of that found in everyday situations (Pearsons et al., 1976; Flexer, 1993; Nabeleck and Nabeleck, 1994). Therefore, S/Ns of 5, 10, and 15 dB were selected for study.

Morales-Garcia and Poole (1972) tested 21 normal-hearing subjects with monosyllabic words and reported 'slightly better' word recognition scores for the right ear than for the left ear at S/Ns of 5 dB and 10 dB. However, the mean differences were only about 3% and were not tested statistically. When words are presented dichotically to normal listeners, slightly better scores have been observed for the right ear than the left ear. This right ear advantage may reflect the dominance of the left hemisphere for speech perception (Studdert-Kennedy and Shankweiler, 1970). However, differential ear effects have not been studied widely for speech-in-noise tests. Therefore, a second purpose of this study was to determine if there is a statistically significant difference in word recognition scores between the right and left ears when normal-hearing subjects are tested with monosyllabic words in a Multitalker Noise at S/Ns of 5, 10, or 15 dB.

A final purpose of this investigation was to assess whether test-retest reliability for monosyllabic words in noise is consistent with the binomial distribution (Thornton and Raffin, 1978; Beattie and Raffin, 1985). This question was addressed with both normal-hearing and hearing-impaired populations.

Experiment 1: Normal-hearing subjects

Method

Subjects

Fifty-one normal-hearing women were tested. They ranged in age from 18 to 30 years, with a mean of 22 years. The subjects reported no history of otoneurologic pathologies. Each subject passed a 15 dB HL (American National Standards Institute, 1989) screening at the octave frequencies from 500 Hz to 4000 Hz (Grason-Stadler, Model GSI 16). Women were chosen because of availability and because Hall (1983) reported no difference in word recognition scores between men and women under 60 years of age. The right and left ears were alternately selected for testing.

Instrumentation and stimuli

The stimuli consisted of the Auditec compact disc recording of the CID W-22 test (lists 1-3, Form A). The Auditec recording of Multitalker Noise (20 voices) served as the competing message. The stimuli were played on a compact disc player (Sharp, Model DX-200) and the output was directed to an audiometer (Grason-Stadler, Model 16) where the stimuli were mixed. The output of the audiometer was directed to an earphone (Telephonic, Model TDH-50P) encased in a cushion (Telephonic, Model P/N-51). The audiometer was calibrated to ANSI-1989 standards using a sound-level meter (Quest Electronics, Model 155) with a one-third octave band filter (Quest Electronics, Model OB-133), coupler (Quest Electronics, Model EC-9A), and microphone (Bruel and Kjaer, Model 4144). The tones preceding the test stimuli and the Multitalker Noise were adjusted to 0 dB on the VU meter, and 0 dB HL was calibrated to 20 dB SPL for both signals. Testing was conducted in an acoustically treated room (Industrial Acoustics Company, Series 400).

In contrast to the monosyllabic words used in this investigation, some authors have argued that sentences are more representative of everyday speech (Jerger and Hayes, 1976; Plomp, 1986). Although Schiavetti et al. (1984) state that sentences are more valid than single words because speech typically is produced in a context with linguistic and acoustic redundancies, the authors present a regression equation that provides an index of contextual intelligibility when only single word scores are available. Moreover, Webster and Snell (1983) comment that classes teaching foreign languages, mathematics, or science often use single word utterances that require audibility of every speech sound. Some authors also advocate single word or syllable tests in order to identify specific sound confusions (Owens and Schubert, 1977; Edgerton and Danhauer, 1979; Dubno and Dirks, 1982).

A Multitalker Noise appears to have reasonable face validity, although there is no convincing evidence that this stimulus is more or less appropriate than other complex stimuli such as cafeteria, industrial, street, or speech noises. The comments of Plomp (1986, pp. 146-147) are pertinent: 'Because the noise radiated from equipment, vehicles, and so forth can be reduced by acoustical treatment, the voices of people surrounding the listener can be considered as the
most prevalent "noise" that has to be accepted as it is. Moreover, speech from a single speaker and voice babble are the most frequent disturbing sounds in everyday situations.

Procedure
The test stimuli were presented at 50 dB HL (70 dB SPL). This level was selected as representative of average conversational speech and because it should elicit scores that approximate PB Max (the highest speech discrimination score attainable). The Multitalker Noise was presented at S/Ns of 5, 10, and 15 dB. Three lists of 50 words each were administered to all subjects (word lists 1A, 2A, and 3A); one list was presented at each S/N. The presentation order of test lists was counterbalanced and the S/Ns were assigned randomly to reduce order effects. Rest periods and verbal encouragement were provided throughout testing in an effort to attain optimal performance. The subjects were provided with response sheets and were required to provide a written response (Nelson and Chaiklin, 1970).

To familiarize the subjects with the procedure, a practice list of 25 words was administered. The practice materials consisted of monosyllabic words in a Multitalker Noise, which were presented at S/Ns of 5, 10, and 15 dB. That is, approximately eight words were presented at each S/N.

Results and discussion
Effects of S/N on word recognition scores
One purpose of this experiment was to establish normal performance (measures of central tendency and variability) on a speech-in-noise test so that these data could serve as a reference for patients with hearing loss. Table 1 presents mean and median word recognition scores from 51 normal-hearing subjects at S/Ns of 5, 10, and 15 dB. The medians agree closely with the means and suggest that either average is a representative measure of central tendency. Table 1 reveals mean scores of about 45% at the 5 dB S/N, 74% at the 10 dB S/N, and 87% at the 15 dB S/N. A one-way analysis-of-variance (ANOVA) for repeated measures revealed a statistically significant difference among S/Ns [F(2, 100) = 3924, P < 0.01]. Tukey's post-hoc test for means (Bruning and Kintz, 1987) revealed significant mean differences among all S/Ns (P < 0.01).

Table 1 also shows similar standard deviations and ranges for all S/Ns; standard deviations varied from 5.3% to 6.2%, and ranges were about 25%.

Table 1. Mean word recognition scores are shown as percentages for normal-hearing subjects as obtained with the compact disc recordings of the CID W-22 words in Multitalker Noise at signal-to-noise ratios of 5 dB, 10 dB, and 15 dB

<table>
<thead>
<tr>
<th>Statistic</th>
<th>5 dB</th>
<th>10 dB</th>
<th>15 dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>45.4</td>
<td>73.9</td>
<td>86.7</td>
</tr>
<tr>
<td>Median</td>
<td>46.0</td>
<td>76.0</td>
<td>86.0</td>
</tr>
<tr>
<td>SD</td>
<td>6.2</td>
<td>5.8</td>
<td>5.3</td>
</tr>
<tr>
<td>Range</td>
<td>24.0</td>
<td>26.0</td>
<td>24.0</td>
</tr>
<tr>
<td>(32–56)</td>
<td>(58–84)</td>
<td>(76–100)</td>
<td></td>
</tr>
</tbody>
</table>

The mean word recognition scores for the normal-hearing subjects are consistent with the findings of Beattie (1989). This author tested 18 normal-hearing subjects with monosyllabic words in a Multitalker Noise and reported mean word recognition scores of 51% at 6 dB S/N, 82% at 12 dB S/N, and 89% at 18 dB S/N. However, our standard deviations of about 6% are smaller than the 8–17% values reported by Beattie (1989).

Ear Differences
Figure 1 presents means and standard deviations for the right and left ears at each S/N. Note that mean scores were similar for both ears at all S/Ns; ear differences were 2.3% at 5 dB S/N, 0.5% at 10 dB S/N, and 0.0% at 15 dB S/N. A two-way ANOVA for mixed designs was conducted. The within subjects factor was S/N with three levels (5, 10, and 15 dB). The between subjects factor was the right ear versus the left ear. Although the ANOVA revealed a statistically significant S/N effect [F(2, 96) = 910, P < 0.01], there was no interaction between the test ear and S/N [F(2, 96) = 0.72, P > 0.01] and no significant ear effect [F(1, 48) = 0.58, P > 0.01].

Contrary to the conclusions of Morales-Garcia and Poole (1972), our data indicate that word recognition scores are equivalent for the right and left ears. That is, no right ear advantage was observed. Therefore, different norms for the two ears do not appear necessary for speech-in-noise tests.
Reliability
The reliability of word recognition scores was assessed by comparing scores on the first-half of each word list (words 1–25) to scores on the second-half of each list (words 26–50). Word recognition score differences were compared to the variability predicted from the binomial theorem at all S/Ns (Thornton and Raffin, 1978). The data revealed that 144 of 153 datum points (94.1%) fell within the predicted 95% confidence interval. These findings are consistent with previous research using monosyllables in both quiet (Thornton and Raffin, 1978; Beattie and Raffin, 1985) and noise (Beattie, 1989) and suggest that the binomial distribution provides a good model for estimating the reliability of monosyllabic word recognition scores.

Experiment 2: Hearing-impaired subjects
Several investigators have commented that monosyllabic words may be too easy to differentiate normal-hearing from sensorineural performance (Carhart, 1965; Penrod, 1994). Moreover, S/Ns that are mildly disruptive for normal hearing subjects can be highly disruptive to hearing-impaired subjects. An S/N should be selected that will separate normal-hearing from hearing-impaired performance, but the S/Ns should not be so difficult that subjects with sensorineural loss cluster around 0%. That is, an S/N should be selected that avoids both floor and ceiling effects. The first experiment, in conjunction with previous research with hearing-impaired subjects, suggested that S/N of 10 dB or 15 dB may be appropriate with subjects having mild-to-moderate hearing loss. Therefore, one purpose of this experiment was to establish word recognition scores on hearing-impaired subjects using monosyllabic words in a background of Multitalker Noise with S/N of 10 dB and 15 dB. We also investigated the test-retest reliability of the monosyllables in noise. Moreover, several illustrative hearing-impaired cases are presented to demonstrate the value of obtaining speech-in-noise word recognition functions.

Method
Most of the procedures described in Experiment 1 were used with the hearing-impaired group. The exceptions are noted below.

Subjects
One ear was tested from 30 subjects with mild-to-moderate sensorineural hearing loss from 250 to 2000 Hz. The subjects ranged in age from 19 to 85 years, with a mean of 66 years. They received a battery of audiologic tests including pure-tone audiometry, speech audiometry, tone decay at 500 Hz, and tympanometry. Word recognition scores were obtained at two or three levels, including the subject's most comfortable listening level and loudness discomfort level (LDL). These scores were obtained by presenting 25 monosyllabic
words via monitored live voice. Pure-tone audiometry revealed interweaving air-conduction and bone-conduction thresholds, word recognition scores did not exhibit rollover, tympanometric peaks were within ±50 daPa (Grason-Stadler, GS1 33), and there was ≤20 dB of tone decay (Olsen and Noffsinger, 1974). The audiologic results are summarized in Table 2.

Mean pure-tone hearing threshold levels (ANSI, 1989) were 25 dB at 500 Hz, 36 dB at 1000 Hz, 46 dB at 2000 Hz, and 64 dB at 4000 Hz. Audiometric configurations ranged from flat to steeply sloping over the 500–4000 Hz range. Maximum word recognition scores (PB Max) ranged from 48% to 100% with a mean of 83%.

Procedure
The test stimuli were presented at the LDL in an effort to elicit scores that would approximate PB Max (Dirks et al., 1981; Beattie and Zipp, 1990). The LDL was obtained using the instructions and procedures suggested by Beattie et al. (1980). Word recognition scores were obtained in quiet and at S/Ns of 10 dB and 15 dB. These stimuli were selected to include listening conditions that would define adequately the word recognition function for subjects with mild-to-moderate hearing loss. Previous research suggested that the selected listening conditions range from relatively easy (quiet) to relatively difficult (10 dB S/N) for subjects with mild-to-moderate hearing loss (Beattie, 1989). The three listening conditions were assigned randomly to reduce order effects. The subjects were instructed to give a verbal response and to guess at every word. The examiner viewed the subjects during testing (Edgerton and Danhauer, 1979). The non-test ear was masked with speech noise which was presented from a portable audiometer.

To familiarize the subjects with the procedure, a practice list of 25 words was administered at their LDL. The practice materials consisted of monosyllabic words in noise, which were presented in quiet and at S/N ratios of 10 dB and 15 dB.

Results
Effects of S/N on word recognition scores
Table 3 presents descriptive statistics for the word recognition scores in quiet and in a background of noise at S/Ns of 10 dB and 15 dB. Table 3 demonstrates that the means and medians are in close agreement. Mean scores were 84.8% in quiet, 59.5% for the 15 dB S/N, and 40.4% for the 10 dB S/N. Standard deviations were about 14%, and the ranges varied from 46% to 60%. A one-way analysis-of-variance revealed a statistically significant difference among the three listening conditions [F(2, 29) = 78, P < 0.01]. Tukey’s post-hoc test (Bruning and Kintz, 1987) indicated significant differences among all means (P < 0.01).

The individual word recognition functions in percent correct are shown in the upper and lower panels of Fig. 2. For comparison, the normal function is presented in each panel as a bold line. The 10 dB and 15 dB S/N means were obtained from the first experiment (74% and 87%) and the mean quiet score (95%) was derived from Beattie, Edgerton, and Svilovec (1977) who also used monosyllabic words. This figure shows that the functions are separated by about 50%, which is expected in view of the variability of pure-tone thresholds and audiometric configurations.

Comparison of word recognition scores

Table 2. Means, medians, standard deviations (SD), and ranges are shown for 240–8000 Hz pure-tone thresholds and loudness discomfort levels in dB HL for subjects with sensorineural hearing loss

<table>
<thead>
<tr>
<th>Frequency in Hz</th>
<th>LDL</th>
<th>PB Max*</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>24</td>
<td>92</td>
</tr>
<tr>
<td>500</td>
<td>25</td>
<td>83%</td>
</tr>
<tr>
<td>1000</td>
<td>36</td>
<td>90</td>
</tr>
<tr>
<td>2000</td>
<td>46</td>
<td>84%</td>
</tr>
<tr>
<td>3000</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>6000</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>8000</td>
<td>77</td>
<td></td>
</tr>
</tbody>
</table>

*PB Max = maximum word recognition scores in percentages
Table 3. Mean word recognition scores as percentages are shown for subjects with mild-to-moderate sensorineural hearing loss as obtained with the Auditec compact disc recordings of the CID W-22 words in quiet and in Multitalker Noise at signal-to-noise ratios (S/N) of 10 dB and 15 dB.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Quiet</th>
<th>15 dB S/N</th>
<th>10 dB S/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>84.8</td>
<td>59.5</td>
<td>40.4</td>
</tr>
<tr>
<td>Median</td>
<td>88.0</td>
<td>62.0</td>
<td>36.0</td>
</tr>
<tr>
<td>SD</td>
<td>12.8</td>
<td>14.0</td>
<td>15.0</td>
</tr>
<tr>
<td>Range</td>
<td>60</td>
<td>52</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>(40-100)</td>
<td>(34-86)</td>
<td>(20-66)</td>
</tr>
</tbody>
</table>

between the normal and hearing-impaired groups indicate that the scores were poorer for the hearing-impaired subjects than for the normal-hearing listeners. This difference was most evident for the speech-in-noise conditions. That is, whereas the two groups differ only by about 10% in the quiet condition, in the noise conditions they differ by about 30%. Although little difference between the quiet and 15 dB S/N conditions was noted for the normal listeners, subjects with mild-to-moderate sensorineural hearing loss exhibited word recognition functions that were about 20% poorer for the 15 dB S/N condition than for the quiet condition. These observations are consistent with previous research indicating that background noise, which has little effect on normal-hearing subjects, can impair substantially the recognition performance of hearing-impaired listeners (Carhart and Tillman, 1970; Beattie, 1989). These findings indicate that subjects with mild-to-moderate sensorineural hearing loss require a more favourable S/N than normal listeners to achieve comparable word recognition scores.

Reliability
The reliability of word recognition scores was assessed by comparing first-half scores (words 1-25) to second-half scores (words 26-50) for the 10 dB and 15 dB S/N conditions. Word recognition score differences were compared to the variability predicted from the binomial theorem. The data revealed that 59 of 60 scores (98%) fell within the 95% confidence interval. These findings are consistent with the normal-hearing data and indicate that the binomial model appears appropriate for estimating the variability of word recognition scores whether they are obtained in quiet or in noise.

The clinician should use a sufficient number of test items in order to reliably identify practical differences among listening conditions. The binomial distribution indicates that word recognition

![Graph](https://via.placeholder.com/150)

Fig. 2. Individual word recognition functions (as percentages correct) for subjects with mild-to-moderate sensorineural hearing loss.

Fifteen subjects each are shown in both the right and left panels. For comparison, the normal function is presented in each panel as a bold line.
scores are dependent on the number of test items. Thus, reliability can be very poor if only a few items are used or excellent if the examiner is willing to administer a few hundred words. In order to identify a difference between two conditions (such as ears, hearing aids, or individuals) of 10%, and assuming an error rate of 5% and word recognition scores in the 20–80% range, approximately 150 test items must be presented in each of the two conditions.

Illustrative examples

Word recognition functions are presented in Figs 3, 4, and 5 from six hearing-impaired subjects to illustrate typical results and to demonstrate the value of testing in noise. Puretone thresholds for each subject are shown at the top of each figure. The bold line in each figure represents mean scores for normal-hearing subjects. Because the standard deviation for normal-hearing subjects was about 6% (Table 1), the 95% confidence interval for normal-hearing performance is ±12%.

Figure 3 presents word recognition functions from three subjects with mild sensorineural hearing loss. These examples show that testing in noise can reveal speech recognition difficulties which are not apparent by testing in quiet. Subject 1 (S1) exhibits excellent word recognition in quiet (96%) despite a mild sensorineural hearing loss in the speech frequencies (PTA = 35 dB) and a moderate-to-severe loss in the higher frequencies. It is evident that measuring speech recognition with monosyllabic words in quiet did not reflect the loss of sensitivity. In contrast, when the monosyllabic words were mixed with Multitalker Noise at S/Ns of 10 dB and 15 dB, word recognition scores were substantially below normal. Compared with the normal mean, the word recognition score was 21% poorer at 15 dB S/N and 40% poorer at 10 dB S/N. In this case, the 10 dB S/N was most useful for revealing the increased difficulty experienced when listening in noise. Subject 2 exhibits hearing within normal limits through 1000 Hz with a sloping hearing loss between 1000 (10 dB HL) Hz and 2000 Hz (40 dB HL). Although the quiet word recognition score suggests little difficulty understanding speech (86%), even the relatively easy 15 dB S/N listening condition shows a much poorer word recogni-

<table>
<thead>
<tr>
<th></th>
<th>600 Hz</th>
<th>1000 Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
<th>8000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>25 dB</td>
<td>30 dB</td>
<td>40 dB</td>
<td>55 dB</td>
<td>105 dB</td>
</tr>
<tr>
<td>S2</td>
<td>35 dB</td>
<td>40 dB</td>
<td>50 dB</td>
<td>70 dB</td>
<td>120 dB</td>
</tr>
<tr>
<td>S3</td>
<td>35 dB</td>
<td>40 dB</td>
<td>45 dB</td>
<td>65 dB</td>
<td>85 dB</td>
</tr>
</tbody>
</table>

Fig. 3. Word recognition functions from three subjects with mild-to-moderate sensorineural hearing loss.

Puretone thresholds are shown at the top of the figure for each subject. The bold line represents mean scores for normal-hearing subjects.

<table>
<thead>
<tr>
<th></th>
<th>600 Hz</th>
<th>1000 Hz</th>
<th>2000 Hz</th>
<th>4000 Hz</th>
<th>8000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4</td>
<td>55 dB</td>
<td>60 dB</td>
<td>65 dB</td>
<td>75 dB</td>
<td>105 dB</td>
</tr>
<tr>
<td>S5</td>
<td>60 dB</td>
<td>65 dB</td>
<td>65 dB</td>
<td>70 dB</td>
<td>85 dB</td>
</tr>
</tbody>
</table>

Fig. 4. Word recognition functions for two subjects with similar hearing losses (moderate, flat to gradually sloping audiograms) but with substantially different scores, particularly for the 15 dB signal-to-noise ratio (S/N).

Puretone thresholds are shown at the top of the figure for each subject. The bold line represents mean scores for normal-hearing subjects.
tion score (34%) than normal-hearing subjects. This finding illustrates that testing in quiet may not reflect the communicative difficulty experienced in noise. Because many listening conditions are characterized by S/Ns <15 dB, these word recognition scores in noise probably are a better indication of the difficulty this individual experiences in everyday situations than the quiet score. Linear regression equations were computed to assess how accurately word recognition scores in noise could be predicted by scores in quiet. These analyses revealed correlation coefficients of 0.74 for the 15 dB S/N condition and 0.66 for the 10 dB S/N condition. However, standard errors of estimate were 10–11%, which indicate that the 95% confidence interval for predicting speech-in-noise scores from speech-in-quiet scores was about ±20%. Thus, these data indicate that speech-in-noise scores cannot be predicted accurately from speech-in-quiet scores. Subject 3 shows a mild hearing loss through 2000 Hz (PTA = 40 dB) and moderate-to-severe loss in the higher frequencies. The word recognition score in quiet was near-normal (86%). It is noteworthy that despite the loss of pure-tone sensitivity, the word recognition scores in noise were only slightly depressed for both the 15 dB S/N (78%) and the 10 dB S/N (62%) conditions. This individual would probably report much better communicative effectiveness in noise than Subject 2.

The subjects in Fig. 4 were selected to show that individuals with similar hearing losses (moderate, flat to gradually sloping audiograms) may exhibit substantially different word recognition functions. That is, these examples suggest that pure-tone thresholds are not accurate predictors of speech recognition ability. Even though Subject 4 had a moderate loss of sensitivity, a high speech recognition score was obtained in quiet (88%) and only slightly reduced scores were observed for the 15 dB S/N (76%) and 10 dB S/N (60%) conditions. In contrast, Subject 5 obtained substantially poorer word recognition scores in noise; a score of only 26% was obtained at 10 dB S/N. These results suggest that there is not a strong relationship between pure-tone thresholds and word recognition scores. To test this assumption, linear regression equations were calculated between various pure-tone thresholds and word recognition scores in quiet and in noise. Table 4 reveals that the highest correlations were found between 2000 Hz and word recognition scores. These correlations were only moderate in magnitude, however, ranging from -0.52 for the 15 dB

Table 4. Pearson product-moment correlation coefficients are shown between pure-tone thresholds (250–8000 Hz) and word recognition scores in quiet and in Multitalker Noise at signal-to-noise ratios (S/Ns) of 10 dB and 15 dB

<table>
<thead>
<tr>
<th>Frequency in Hz</th>
<th>Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>250</td>
</tr>
<tr>
<td>Quiet</td>
<td>-0.04</td>
</tr>
<tr>
<td>15 dB S/N</td>
<td>0.13</td>
</tr>
<tr>
<td>10 dB S/N</td>
<td>0.08</td>
</tr>
</tbody>
</table>
S/N condition to -0.68 for the 10 dB S/N condition. Moreover, large standard errors of estimate of 11–12% were observed; this suggests that the 95% confidence interval for predicting the word recognition score from the 2000 Hz threshold is ±24%. The word recognition scores in Fig. 4 also illustrate that testing in quiet may not reflect the relative difficulty an individual has in noise as compared with normal subjects. That is, speech-in-noise scores may enable the clinician to provide a better estimate of communicative ability than if based only on speech-in-quiet scores. These data also suggest that compensation decisions should not be based only on puretone thresholds and speech recognition testing in quiet.

Figure 5 presents a word recognition function from Subject 6 who exhibits a moderate hearing loss across the speech frequencies. This subject is of interest because the quiet word recognition score (40%) clearly differentiated this subject from normal performance. Although very poor scores also were obtained at the 10 dB and 15 dB S/N conditions, testing in noise was not necessary to differentiate this patient from normal hearing subjects. That is, although monosyllabic words in quiet may be too easy to demonstrate communicative difficulty with some patients having mild-to-moderate hearing loss, other patients may exhibit considerable difficulty when compared to individuals with normal hearing.

Conclusions
The first experiment described normal performance for monosyllables in noise so that these data could serve as a reference for patients with hearing loss. The second experiment was designed to assess how subjects with mild-to-moderate sensorineural hearing loss perform in comparison to normal listeners and to gather preliminary data on the contribution of speech-in-noise testing to the management of the hearing-impaired listener.

The data provide further evidence suggesting that speech-in-noise tests can enable a better understanding of communicative difficulties and can aid in the management of hearing-impaired persons. Subjects with mild-to-moderate sensorineural hearing loss exhibited poorer word recognition scores in noise than the normal-hearing subjects. These results suggest that background noise which has little effect on normal-hearing subjects can substantially affect the word recognition performance of hearing-impaired listeners.

Both the normal-hearing and hearing-impaired test-retest data indicate that the binomial model is appropriate for estimating the variability of word recognition scores when they are obtained in noise. These findings are consistent with speech-in-quiet data.

The literature indicates that speech-in-noise testing may help assess communicative function, select an ear for amplification, award compensation, differentiate among hearing aids, identify site-of-lesion, or to assess the effects of medical or non-medical treatment. Although monosyllabic words in noise were used in this investigation, no single speech-in-noise test is appropriate for all questions and patients. Instead, a battery of speech-in-noise tests may be selected to address those questions of interest. Regardless of the particular tests selected, speech-in-noise testing can provide insight into communicative difficulties experience by the hearing-impaired and can help assess the effectiveness of remediation.

References
Beattie RC, Svihovec DA, Carmen RE, Kunkel HA. Loudness discomfort level for speech:
Gatehouse S, Haggard MP. The effects of air-bone gap and presentation level on word identification. Ear Hear 1987; 8: 140–146.
Olsen WO, Noffsinger D. Comparison of one