J Am Acad Audiol 19:507–518 (2008)

Predicting Word-Recognition Performance in Noise by Young Listeners with Normal Hearing Using Acoustic, Phonetic, and Lexical Variables

DOI: 10.3766/jaaa.19.6.6

Rachel McArdle*
Richard H. Wilson†

Abstract

Purpose: To analyze the 50% correct recognition data that were from the Wilson et al (this issue) study and that were obtained from 24 listeners with normal hearing; also to examine whether acoustic, phonetic, or lexical variables can predict recognition performance for monosyllabic words presented in speech-spectrum noise.

Research Design: The specific variables are as follows: (a) acoustic variables (i.e., effective root-mean-square sound pressure level, duration), (b) phonetic variables (i.e., consonant features such as manner, place, and voicing for initial and final phonemes; vowel phonemes), and (c) lexical variables (i.e., word frequency, word familiarity, neighborhood density, neighborhood frequency).

Data Collection and Analysis: The descriptive, correlational study will examine the influence of acoustic, phonetic, and lexical variables on speech recognition in noise performance.

Results: Regression analysis demonstrated that 45% of the variance in the 50% point was accounted for by acoustic and phonetic variables whereas only 3% of the variance was accounted for by lexical variables. These findings suggest that monosyllabic word-recognition-in-noise is more dependent on bottom-up processing than on top-down processing.

Conclusions: The results suggest that when speech-in-noise testing is used in a pre- and post-hearing-aid-fitting format, the use of monosyllabic words may be sensitive to changes in audibility resulting from amplification.

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

Abbreviations: BKB = Bamford-Kowal-Bench; CID W-22 = Central Institute for the Deaf W-22 word lists; CV = consonant–vowel; HHH = High word frequency, High neighborhood density, and High neighborhood frequency; HHL = High word frequency, High neighborhood density, and Low neighborhood frequency; HLH = High word frequency, Low neighborhood density, and High neighborhood frequency; HLL = High word frequency, Low neighborhood density, and Low neighborhood frequency; IEEE = Institute of Electrical and Electronics Engineers; LHH = Low word frequency, High neighborhood density, and High neighborhood frequency; LHL = Low word frequency, High neighborhood density, and Low neighborhood frequency; LLH = Low word frequency, Low neighborhood density, and High neighborhood frequency; LLL = Low word frequency, Low neighborhood density, and Low neighborhood frequency; NAM = Neighborhood Activation Model; PB-50 = phonetically balanced 50 (word list); rms = root-mean-square; S/N = SNR = signal-to-noise ratio; SSN = speech-spectrum noise.

*Bay Pines VA Healthcare System, Bay Pines, FL, and University of South Florida, Tampa, FL; †VA Medical Center, Mountain Home, TN, and East Tennessee State University, Johnson City, TN

Rachel A. McArdle, Ph.D., Bay Pines VA HCS, Audiology (126), PO Box 5005, Bay Pines, FL 33744; Phone: (727) 398-9395; Fax: (727) 319-1209; E-mail: Rachel.mcardle@va.gov

Portions of this paper were presented at the American Academy of Audiology Convention, March 2006, Minneapolis, MN, and at the International Symposium on Auditory and Audiological Research (ISAAR), August 2007, Helsingør, Denmark. An abbreviated version of the manuscript was included in the ISAAR proceedings.

Sumario

Propósito: Analizar los datos de reconocimiento correcto del 50% que surgieron del estudio de Wilson y col. (este ejemplar) y que fueron obtenidos de 24 sujetos con audición normal; también examinar si las variables acústicas, fonéticas o léxicas pueden predecir el desempeño en el reconocimiento para palabras monosilábicas presentadas en ruido dentro del espectro del lenguaje.

Diseño de Investigación: Las variables específicas son: (a) variables acústicas (p.e., nivel efectivo de presión sonora de raíz media cuadrada, duración), (b) variables fonéticas (p.e., rasgos de consonantes tales como modo, lugar y sonorización para los fonemas iniciales y finales; fonemas vocálicos), y (c) variables de léxico (p.e., frecuencia de la palabra, familiaridad de la palabra, densidad del vecindad, frecuencia del vecindad).

Recolección y Análisis de los Datos: El estudio descriptivo y de correlación examinará la influencia de las variables acústicas, fonéticas y léxicas sobre el desempeño en el reconocimiento del lenguaje en medio de ruido.

Resultados: Los análisis de regresión demostraron que el 45% de la variancia en el punto 50% fue producto de las variables acústicas y fonéticas, mientras que sólo el 3% de la variancia fue debida a las variables léxicas. Estos hallazgos sugieren que el reconocimiento de palabras monosilábicas en ruido es más dependiente de un procesamiento de abajo hacia arriba que de arriba hacia abajo.

Conclusiones: Los resultados sugieren que cuando se usa la evaluación de lenguaje en ruido en un formato pre- y post-adaptación del auxiliar auditivo, el uso de palabras monosilábicas puede ser sensible a los cambios en la audibilidad resultantes de la amplificación.

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

Abreviaturas: BKB = Bamford-Kowal-Bench; CID W-22 = lista de palabras W-22 del Instituto Central de Sordo; CV = consonante vocal; HHH = alta frecuencia de la palabra, alta densidad de vecindad y alta frecuencia de vecindad; HHL = alta frecuencia de la palabra, alta densidad de vecindad y baja frecuencia de vecindad; HLH = alta frecuencia de la palabra, baja densidad de vecindad y alta frecuencia de vecindad; HLL = alta frecuencia de palabra, baja densidad de vecindad y baja frecuencia de vecindad; IEEE = Instituto de Ingenieros Eléctricos y Electrónico; LHH = baja frecuencia de la palabra, alta densidad de vecindad y alta frecuencia de vecindad; LHL = baja frecuencia de la palabra, alta densidad de vecindad y baja frecuencia de vecindad; LLH = baja frecuencia de la palabra, baja densidad de vecindad y alta frecuencia de vecindad; LLL = baja frecuencia de la palabra, baja densidad de vecindad y baja frecuencia de vecindad; NAM = Modelo de Activación de Vecindad; PB-50 = lista de palabras fonéticamente balanceadas 50; rms = raíz media cuadrada; S/N = SNR, tasa señal-ruido; SSN = ruido de especto del lenguaje

INTRODUCTION

Speech-in-Noise Testing

Several speech-in-noise tests have been developed for research and clinical use; the majority of them use sentence materials (e.g., the Connected Speech test, Cox et al, 1987; the Hearing in Noise Test, Nilsson et al, 1994; the QuickSIN, Killion et al, 2004; and the BKB-SIN, Etymotic Research, 2005). Although sentence materials have good face validity because they are more reflective of "real world" listening than are isolated word materials, performance on sentence measures can be influenced by factors other than audibility, such as syntactic and semantic context. The influence of context on performance can be representative of how an individual communicates in everyday life, which is a top-down process, but it may not reflect basic auditory function, which is a bottom-up process.

One of the roles of clinical speech-in-noise testing is to help optimize the fitting of sensory aids (e.g., hearing aids and cochlear implants). With that purpose, then, one would want to minimize the top-down influences on performance. As demonstrated by Miller (1951) and recently by our laboratories (Wilson et al, this issue) the goal might be better accomplished with isolated words than with sentence materials, which, as previously indicated, contain syntactic and semantic cues for recognition.

Boothroyd and Nittrouer (1988) suggested that the influence of linguistic context on word recognition could be explained by a j-factor. The j-factor is the ratio of the logarithms for the recognition probability of the whole word and the recognition probability for the parts within the whole word. Boothroyd and Nittrouer showed that j-factors for nonsense words equaled 3.0, suggesting that each phoneme was being perceived independently, whereas for real words the j-factors were ~2.4–2.6, suggesting that words have some lexical context.

Although linguistic context has been shown to influence recognition performance for monosyllabic words, the addition of syntactic and semantic context (i.e., meaningful sentences) provides an even greater increase in recognition performance. A recent study conducted by the authors (Wilson et al, this issue) examined the recognition performance for sentences with rich semantic and syntactic context (Bamford-Kowal-Bench, or BKB, sentences), sentences with syntactic context and slight semantic context (Institute of Electrical and Electronics Engineers, or IEEE, sentences), and monosyllabic words (NU–6). The results demonstrated an inverse relationship between context and recognition performance. As context decreased, the signal-to-noise ratio (SNR, S/N) for the 50% points on the recognition functions increased. Thus, monosyllables required a higher SNR than did either of the sentence materials, suggesting that monosyllabic words may not be representative of everyday conversation but are more reflective of basic auditory perception.

Bottom-up and Top-down Processing

Historically, speech-perception researchers have contended that the process of recognizing speech is organized hierarchically, with the acoustic properties of an utterance as the foundation. The hierarchy view suggests that bottom-up processing refers to the use of acoustic information at the "bottom" of the hierarchy (i.e., audibility) to contribute to the recognition of an utterance, whereas top-down processing refers to the use of higher-level cognitive information found toward the "top" of the hierarchy.

Both word and sentence materials activate bottom-up and top-down processing during a recognition task. Bottom-up and top-down processing are inversely related for word and sentence materials. Word recognition uses bottom-up processing more than sentence recognition does, whereas sentence recognition uses top-down processing more than word recognition does. Although the use of monosyllabic words is thought to reduce the influence of higher level processing such as semantic and syntactic contextual cues, past studies, including those examining the Neighborhood Activation Model (NAM), suggest that recognition performance for isolated words is also influenced by top-down lexical processing (Luce and Pisoni, 1998).

The NAM is a theory that describes the organization of perceptual information in long-term memory and the way in which that information is used for lexical access. The model assumes that when a spoken word is perceived, it activates a set of representations of similar-sounding words in memory. This set of phonetically similar words is called a *similarity neighborhood*. A similarity neighborhood consists of a target word and all other words that can be created from the target word by adding, deleting, or substituting a single phoneme. For example, the neighborhood for the word *cat* would include the words *bat*, *at*, *scat*, *cut*, and *kit* (and all other words differing from *cat* by a single phoneme).

According to the NAM, the primary organizational units of phonetic information in long-term memory are the similarity neighborhoods. The model predicts that the acoustic information (stimulus input, bottom-up) of a word is received and activates all of the corresponding acoustic–phonetic patterns in long-term memory that match the incoming acoustic information. If the acoustic–phonetic pattern of the word corresponds to an actual word in long-term memory, then a word-decision unit is activated for each corresponding word, activating a "neighborhood" of words with similar acoustic–phonetic structure (i.e., bottom-up processing). Once the neighborhood is activated, the word-decision units compete with each other during the selection phase.

Top-down processing adjusts the activation level of the individual word-decision units by biasing the system toward words that have a high frequency of occurrence in the English language. Bottom-up processing continues to affect the activation levels so that as more auditory input (information) is received, the activation levels of the word decision units that no longer match the acoustic input are lowered, dropping them out of the competition (Luce, 1986; Luce et al, 1990). Word recognition or selection occurs when the bottom-up and top-down processing converge on the best match from the activated neighborhood (Luce and Pisoni, 1998).

In a series of experiments with young adults who had normal hearing, Luce and his colleagues (Luce, 1986; Cluff and Luce, 1990; Luce et al, 1990) demonstrated that the difficulty of isolating a target word from its neighbors is determined by word frequency, neighborhood density, and neighborhood frequency. *Word frequency* is a value that refers to the number of times a word is used in the language. Previous studies (Sommers, 1996; Luce and Pisoni, 1998; Dirks et al, 2001) demonstrate that low-frequency (of occurrence) words are more difficult to recognize than high-frequency (of occurrence) words because of lower resting activation levels.

Likewise, words in a high-frequency neighborhood are more difficult to recognize than words in a low-frequency neighborhood because of the higher levels of competition within the neighborhood. *Neighborhood density* refers to the absolute number of words occurring in any given similarity neighborhood. A high-density word has numerous lexical neighbors, making it sound similar to many other words (e.g., *tot*), whereas a low-density word has very few neighbors and sounds like very few other words (e.g., *golf*). Accordingly, words from low-density neighborhoods are more easily recognized than are words from high-density neighborhoods. *Neighborhood frequency* refers to the average frequency of words in a lexical neighborhood (i.e., the mean frequency of occurrence in the English language of all words in a similarity neighborhood). Words in low-frequency neighborhoods are recognized more accurately than are words in high-frequency neighborhoods.

These three factors—word frequency, neighborhood density, and neighborhood frequency—can be used to characterize the lexical properties of monosyllabic words. Studies that have examined the effects of lexical difficulty on monaural word-recognition performance typically compared performance between lexically EASY words and lexically HARD words. The definition of EASY and HARD words, however, varies among investigations. In studies in which all three factors are considered (Luce and Pisoni, 1998; Dirks et al, 2001), a lexically EASY word is defined as a high-frequency word from a low-density, low-frequency neighborhood.

At the other extreme, a lexically HARD word is defined as a low-frequency word from a high-density, high-frequency neighborhood. In other studies (Sommers, 1996; Sommers et al, 1997; Sommers and Danielson, 1999), word frequency is controlled, and EASY and HARD words are defined by their neighborhood-density and neighborhood-frequency characteristics. In this context, an EASY word is a word from a low-density, low-frequency neighborhood, whereas a HARD word is a word from a high-density, high-frequency neighborhood. Alternatively, Kirk et al (1997) used word frequency and neighborhood density to define EASY and HARD words, whereas Bell and Wilson (2001) used only neighborhood density and neighborhood frequency without considering the individual target-word frequency values.

Acoustic, Phonetic, and Lexical Variables That Influence Recognition

Over the past 50 years the effects of several variables on speech intelligibility have been studied. Two acoustic variables that have been shown to influence the recognition performance of words are the effective root mean square sound pressure level and the duration of the word. Acoustic variables are those variables that define the physical principles of sound waves. There is a direct relation between rms and recognition performance (Egan, 1948). For a given word, a higher rms level translates to a higher presentation level, which, in turn, increases audibility. As the audibility of a word increases, recognition performance increases.

There is also a direct relation between the duration and intelligibility of words. Egan (1948) presented five groups of 20 words that varied in number of sounds from three to six sounds per word to a group of listeners and observed that recognition performance improved markedly as the number of sounds per word increased. Similarly, Silverman and Hirsh (1955) indicated that the Rush Hughes recordings of the PB-50 (phonetically balanced 50) words produced poorer performance than did the Hirsh (or Reynolds) recordings of the same words because the words produced by the former speaker were shorter (i.e., at a faster rate) than the words produced by the latter speaker.

Phonetic variables are the categorical variables that define the actual properties of individual speech sounds. Phonetic variables thought to influence recognition performance are consonant features and the vowel phoneme (Mason, 1945; Black, 1952; Walden and Montgomery, 1975; Dubno et al, 1982; Gordon-Salant, 1985). Early work in this area examined the influence of both phonemes and vowels on recognition ability. Black (1952) extended the work of Mason (1945) and presented 3,697 monosyllables to 160 panels of 9–12 Navy personnel. Black found that the phonetic content of consonants significantly influenced the relative difficulty of word-recognition performance.

In terms of vowel phonemes, Mason (1945) found the use of diphthongs to influence recognition performance; however, Black (1952) examined only short and long vowels and found no difference in performance for words containing either vowel type. More recently, the findings regarding the influence of different vowel phonemes on recognition in noise are inconsistent. Dubno et al (1982) and Dubno and Levitt (1981) found that consonants followed by /a/ were recognized at a higher accuracy than were consonants followed by /i/ and /u/. The reverse was reported by Gordon-Salant (1985), who reported higher accuracy was found for consonants following /i/ and /u/ than for /a/. A third set of findings was reported by Wang and Bilger (1973), who observed that recognition accuracy was highest for consonants following /u/ and poorest for consonants following /i/. Methodological issues (e.g., level differences, noise differences) among these studies may have led to the conflicting results.

The influence of lexical variables on the relative ease or difficulty of recognition performance also has been studied. Most studies of nonacoustic variables have focused on word frequency, which refers to the number of times a word occurs in the language (e.g., Howes, 1957; O'Neill, 1957; Rosenzweig and Postman, 1957). As mentioned previously, word frequency is one of the fundamental variables discussed in the NAM. Historically, word-frequency values are based on the data collected by Kucera and Francis (1967), who sampled a million words in printed text to provide frequency counts of words per million. For example, a high-frequency word such as *was* has a frequency count of 9,816 per million, whereas a low-frequency word such as *cat* has a frequency count of 23 per million. Although the precise mechanism is not clear, research results demonstrate processing advantages for high-frequency words (Soloman and Postman, 1952; Rosenzweig and Postman, 1957; Broadbent, 1967). Word frequency biases the retrieval process toward choosing high-frequency words over low-frequency words.

Similar to word frequency is the lexical variable of word familiarity (e.g., Black, 1952). Subjectively rated word familiarity has been reported as a more accurate index of word experience than objective counts of words in texts (Gernsbacher, 1984). Although there is a debate as to the validity of word familiarity as an estimate of word-exposure rate, word familiarity has been found to be an effective predictor of word recognition (Auer et al, 2000). As previously mentioned, Black examined the relationship between the recognition performances of almost 4,000 words and variables that may influence recognition performance. Black found that even among common words, word familiarity influenced recognition performance so that more-familiar words are recognized more accurately.

Most models of speech recognition address the effect of word frequency and lexical competition; however, the lexical variables of neighborhood density and neighborhood frequency are specific to the NAM (Luce and Pisoni, 1998). As described previously, neighborhood density and neighborhood frequency have been cited as variables that can influence the relative ease or difficulty with which a word is recognized (e.g., Dirks et al, 2001).

The aforementioned acoustic and phonetic variables are associated with the bottom-up processing that occurs during a word-recognition task, whereas the lexical variables refer to the top-down processing. Examination of the relative importance of bottom-up and top-down processing for monosyllabic words would require a large set of words spoken by the same speaker in order to minimize or eliminate the confounding variable of speaker differences (Kreul et al, 1969).

A previous study by the authors was conducted on 24 young listeners with normal hearing to determine the psychometric properties of 12 common monosyllabic word lists presented in speech-spectrum noise (SSN) at four SNRs (Wilson et al, this issue). The monosyllables were recorded by a female speaker from the following word-lists: PB-50, Lists 8–11 (Egan, 1948); CID W-22 (Central Institute for the Deaf W-22 word lists), Lists 1–4 (Hirsh et al, 1952); and NU–6, Lists 1–4 (Tillman and Carhart, 1966). Although these monosyllabic word lists are more commonly presented in quiet, speech-spectrum noise was used to homogenize the intersubject audibility and to reduce ceiling effects. The metrics of interest were the mean percentage correct for each word list at four SNRs and the SNR at which the mean 50% point occurred. The results showed almost identical psychometric functions for the percentage of correct data obtained for the PB-50, CID W-22, and NU–6 word lists with <0.5 dB differences at the 50% points among the three sets of materials. These findings indicate that performance differences among the various materials reported in the past literature probably reflected speaker differences more than differences in difficulty among the tokens of the three sets of test materials.

In addition to examining list differences, Wilson et al (this issue) also examined the mean psychometric functions of each of the 546 monosyllables and found that the 50% points ranged 17.9 dB, from −8.3 dB S/N (*choice*) to 9.7 dB S/N (*bob*). Given this wide range of recognition performances among monosyllables, it was of interest to identify the variables that affected the relative ease or difficulty with which the monosyllable words were recognized.

For the current study, the Wilson et al (this issue) data were analyzed to identify a set of variables that were predictive of the 50% point of recognition performance. The following types of variables were included in the evaluation: (a) acoustic variables (i.e., rms and duration), (b) phonetic variables (i.e., consonant features—such as initial and final consonant manner, place, and voicing—and vowel phoneme), and (c) lexical variables (i.e., word frequency, word familiarity, neighborhood density, and neighborhood frequency). Although many of these variables have been examined previously, the relative importance of each variable on word-recognition-in-noise performance has not been examined.

METHODS

Psychometric Data

The Wilson et al (this issue) study details the materials, participants, and procedures used to generate the psychometric data on the 546 monosyllabic words. Briefly, the data were obtained from 24 young listeners with normal hearing. Words from four lists from each of the PB-50, CID W-22, and NU–6 materials were interspersed and recorded by a female speaker, were edited, and were presented at four SNRs. The level of the SSN was fixed at 72-dB SPL with the level of the speech varied from 65-dB SPL (−7 dB S/N) to 80-dB SPL (8-dB S/N) in 5-dB steps. Thus, each of the 546 words was presented to each listener at each of the four SNRs using a randomized design. The 50% point on the psychometric function for each word was established with the Spearman-Kärber equation (Finney, 1952; Wilson et al, 1973).

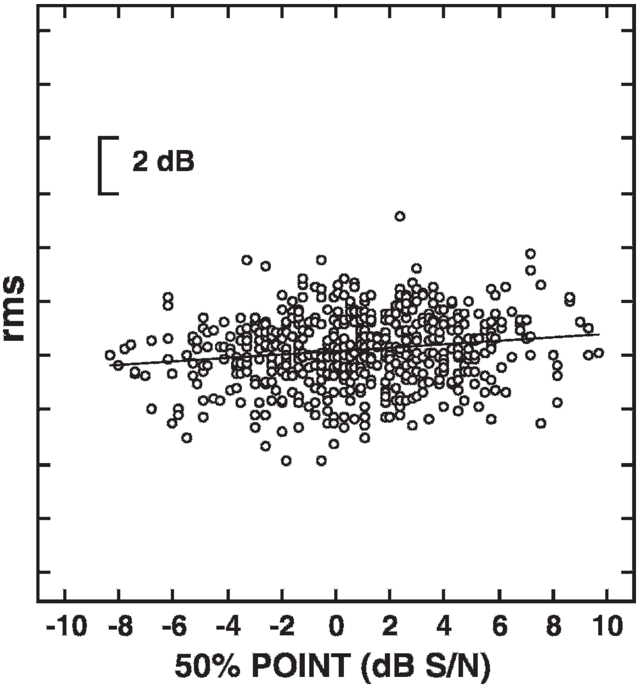

Figure 1. A bivariate plot of the mean 50% points (abscissa) and root-mean-square (rms; ordinate) of each of the 490 monosyllabic words. The regression line is shown with a slope of 0.06 dB/dB.

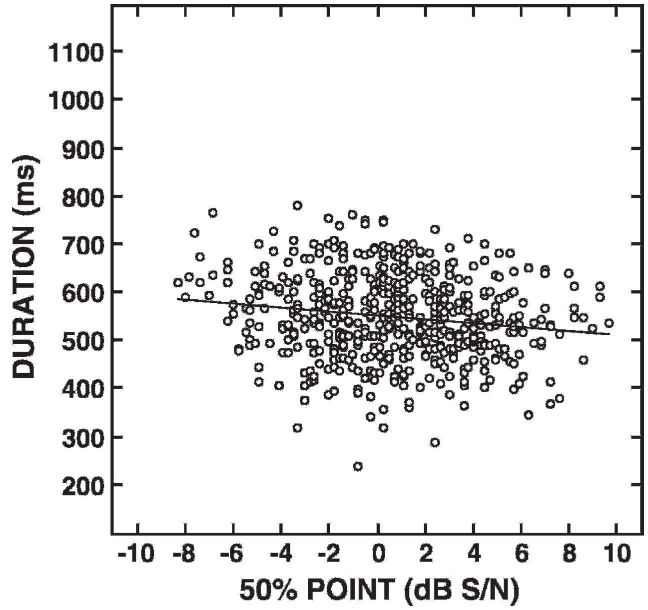

Figure 2. A bivariate plot of the mean 50% points (abscissa) and duration (ordinate) of each of the 490 monosyllabic words. The linear regression has a slope of −4.1 msec/dB.

Procedures

The rms values (dB, re: maximum digitization range) and the duration values (msec) of each of the monosyllabic words were measured using sound-editing software (Adobe Audition 2.0). Additionally, the articulatory characteristics—including manner, place, and voicing of the initial and final phonemes and the vowels—were transcribed from the recorded materials by a phonologist. The 20,000-word Hoosier Mental Lexicon (Nusbaum et al, 1984) was used to generate the values of each of the lexical variables (i.e., word frequency, familiarity, neighborhood density, and neighborhood frequency) for each of the monosyllables that were available in the lexicon. The initial 546[1] monosyllabic words reported by Wilson et al (this issue) included 56 monosyllabic words that were not listed in the Hoosier Mental Lexicon, so the corpus of words used in the current analysis was reduced to 490.

As in previous studies (e.g., Sommers, 1996; Dirks et al, 2001), median splits were performed for word frequency, neighborhood density, and neighborhood frequency in order to categorize the words as high (H) or low (L) with respect to the median. Values for each of the three variables that were above the median were coded as H, whereas values below the median were coded as L. Each target word was coded as H or L for all three variables. Additionally, the absolute values provided by the Hoosier Lexicon for all three lexical factors were included in the data set. (A table with the lexical data for the 490 words is available from the authors.)

RESULTS AND DISCUSSION

Acoustic Variables

The rms values of the whole word are plotted in Figure 1 as a function of the recognition performances at the 50% points (in dB S/N) for each of the 490 words. A Pearson Product-Moment correlation was significant ($r = 0.16$, $p < .01$), but the size of the correlation suggests a weak relationship between rms and recognition performance. The linear regression used to describe the data had a slope of 0.06 dB/dB, which approximates a slope of zero. In all probability the slope is more indicative of the relation between rms and recognition performance, with the finding of statistical significance attributable to the large number of samples.

In Figure 2, word duration is plotted as a function of the recognition performances at the 50% point. Again, a Pearson Product-Moment correlation was significant ($r = -0.18$, $p < .01$), but the size of the correlation suggests a weak relationship between duration and recognition performance. The relationship shows that shorter words required a more positive SNR than the longer words at the 50% point. The slope of the linear regression used to describe the data in Figure 2 was −4.1 msec/dB, which with consideration to the magnitude of the duration metric is a very gradual slope that is not considered particularly noteworthy.

Table 1. Mean 50% Points (dB S/N) and Standard Deviations (dB) for the Consonant Features of Manner, Place, and Voicing of the Initial and Final Phonemes

	Phoneme Position							
	Initial	Final	Initial	Final	Initial	Final	Initial	Final
Manner	Stops		Fricatives		Affricates		Nasals	
Mean	1.0	0.8	−0.1	−0.6	−1.9	−2.9	0.9	1.8
SD	3.3	3.5	3.3	3.0	4.5	2.9	3.2	2.8
	Liquids		Glides					
Mean	2.8	2.5	0.6	—				
SD	2.6	3.0	2.3	—				
Place	Bilabials		Labiodental		Dental		Alveolar	
Mean	2.4	2.2	0.3	0.8	1.6	−0.5	0.7	0.6
SD	3.2	3.2	2.6	2.9	2.3	2.6	3.5	3.4
	Alveopalatal		Velar		Labio-velar		Glottal	
Mean	−1.7	−2.6	0.7	1.1	0.8	—	0.9	—
SD	3.8	2.6	3.0	3.4	2.4	—	3.3	—
Voicing	Voiceless		Voiced					
Mean	−0.2	−0.7	1.6	1.9				
SD	3.3	3.1	3.2	3.2				

As with the rms data, the finding of statistical significance for duration most likely is attributable to the large data set. The finding of only a weak relationship between the recognition performance at the 50% point and each acoustic variable (i.e., rms and duration) was somewhat expected, given that the stimuli were recorded at similar levels in conjunction with a carrier phrase and were restricted to monosyllables to reduce interstimulus variability for the initial experiment (Wilson et al, this issue). The range in terms of duration for the monosyllabic words, however, was wide in that the longest word (750 msec) was almost three times the length of the shortest word (240 msec). The same was true for rms with the weakest word (−25.8 dB, re: maximum digitization range) 9 dB below the strongest word (−16.8 dB).

Table 1 lists the mean 50% correct recognition points (and standard deviations) for the individual words as a function of consonant features such as manner, place, and voicing of the initial and final phoneme speech sounds. Consider first the *Manner* section of Table 1. The 50% points ranged from the most difficult to understand in noise, liquids (i.e., /l,r/)—with mean 50% points of 2.8-dB S/N and 2.5-dB S/N for the initial and final positions, respectively—to the easiest to understand in noise, affricates (i.e., /dʒ, tʃ/)—with mean 50% points of −1.9-dB S/N and −2.9-dB S/N for the initial and final positions, respectively. For the *Manner* section of Table 1, the decibel range in terms of 50% points from the easiest to the hardest category was 4.7 dB and 5.4 dB for initial and final position, respectively.

As shown in the *Place* section of Table 1, bilabials (i.e., /b,p,m/) were the most difficult to understand in noise with mean 50% points of 2.4-dB S/N and 2.2-dB S/N for the initial and final positions, respectively, whereas the alveopalatals (i.e., /dʒ, tʃ, ʃ, ʒ/) were the easiest to understand with mean 50% points of −1.7-dB S/N and −2.6-dB S/N for initial and final positions, respectively. For the *Place* section in Table 1, the decibel ranges in terms of 50% points were similar to those of the *Manner* section, with a decibel range between the easiest and most difficult *place* category of 4.1 dB and 4.8 dB for initial and final position, respectively.

Interestingly, the phonemes (i.e., /dʒ, tʃ/) belonging to the affricate *manner* category and the alveopalatal *place* category were the easiest categories to recognize in noise for both initial and final position. Thus, it is unclear as to whether it is the consonant feature of manner or place that positively affected recognition performance. Previous studies have also found manner and place to affect recognition performance significantly when using nonsense consonant–vowel (CV) combinations to create consonant confusion matrices (Walden and Montgomery, 1975; Bilger and Wang, 1976; Dubno et al, 1982). Similarly, the results of the current study support previous reports indicating that the phoneme in the initial position of an utterance was easier to recognize than the phoneme in the final position of the utterance (e.g., Dubno et al, 1982).

The final consonant feature highlighted in Table 1 is that of voicing. The voiceless phonemes in the initial (−0.2-dB S/N) and final (−0.7-dB S/N) positions were recognized at lower SNRs than were the voiced phonemes in the initial (1.6-dB S/N) and final (1.9-dB S/N) positions. This finding is consistent with that of Miller and Nicely (1955), who reported that voiceless phonemes in English are more intense, given that voiceless phonemes are aperiodic and noisier than voiced phonemes. However, other investigators (Wang and Bilger, 1973; Walden and Montgomery, 1975) subsequently reported that voicing was not a dominant feature during perception experiments in quiet, although Walden and Montgomery did state that their finding may reflect different testing paradigms. Walden and Montgomery used suprathreshold similarity judgments in quiet, where consonant level is not as important as it is for a recognition task in noise, such as the one used in the current study as well as in Miller and Nicely's study.

The data in Table 2 show the number of specific vowel phonemes included in the sample, the mean 50% points in SNR, and the standard deviations for 18 vowel phonemes. The vowel phonemes are listed from easiest to understand in noise (/ju/, −2.2-dB S/N) to most difficult to understand in noise (/ɑ/, 1.9-dB S/N), representing a 4.1-dB range. A one-way analysis of variance (ANOVA) found no significant difference, $F(17, 489) = 1.45$, $p > .05$, in mean 50% points among the 18 vowel-phoneme categories.

Previous studies that examined the effect of vowel phonemes used nonsense CVs with /ɑ, i, u/ as the target vowel phonemes (e.g., Wang and Bilger, 1973; Gordon-Salant, 1985, Dubno and Levitt, 1981; Dubno et al, 1982). As previously mentioned, the results from past studies were not straightforward. Specifically, for the three vowel phonemes used in the earlier studies (i.e., /ɑ, i, u/), recognition performances in the current study were easiest for words with the vowel phoneme /u/, followed by the vowel phoneme /i/, and were most difficult for words with the vowel phoneme /ɑ/. This finding is similar to that of Gordon-Salant (1985), who found higher accuracy values for consonants following /i/ and /u/ than for /a/.

Lexical Variables

The words were categorized as H or L for word frequency, neighborhood density, and neighborhood frequency, so that words in the HLL group had *H*igh word frequency, *L*ow neighborhood density, and *L*ow neighborhood frequency. Table 3 shows the eight NAM categories in the first column, the number of words per category in the second column, and the percentage of words for each category represented in the data set in the third column. Although the total number of words for each of the eight lexical categories varies widely, ranging from 35 to 96 words, by using the values from the middle column in Table 3, one will find that the following

Table 2. The Mean 50% Correct Recognition Points (dB S/N) and Standard Deviations (dB) for the Words That Contained the Respective Vowel Phonemes

Vowel Symbol	Examples	Number of Words	50% pt	*SD*
/ju/	cued, you, juice	9	−2.2	3.2
/ɛr/	hair, there, tare	4	−1.3	2.8
/aʊ/	how, shall, doubt	13	−0.8	4.4
/ɝ/	hurt, search, third	18	−0.2	2.9
/ɑr/	hard, carve, bar	16	−0.1	4.0
/oʊ/	code, foe, though	34	0.0	4.6
/u/	hoot, goose, chew	22	0.2	3.1
/ae/	had, bath, and	60	0.3	3.6
/ɛ/	head, chess, said	35	0.3	3.2
/eɪ/	hate, chain, drake	39	0.4	3.4
/ɔɪ/	boyd, void, toy	6	0.6	3.0
/ʊ/	hood, cook, should	11	1.1	2.1
/ʌ/	bud, tub, scrub	47	1.1	3.8
/i/	heat, keen, seize	34	1.2	3.4
/aɪ/	high, rhyme, wife	35	1.2	3.2
/ɔ/, /ɔr/	jaw, flaunt, hog	40	1.4	3.4
/ɪ/, /ɪr/	hit, bliss, gin	61	1.4	3.0
/ɑ/	hot, cod, wash	6	1.9	4.1

similarities for each lexical characteristic (word frequency, neighborhood density, and neighborhood frequency) were found: (a) half of the sample ($n = 244$) are high-frequency words and the other half ($n = 246$) are low-frequency words, (b) half of the words are from a high-density neighborhood ($n = 240$) and the other half from a low-density neighborhood ($n = 250$), and (c) half of the words are associated with a high-frequency neighborhood ($n = 247$) and the other half with a low-frequency neighborhood ($n = 243$). The symmetry for high and low classifications found for each of the lexical categories suggests that each lexical characteristic had equal representation in the 490-word sample.

Figure 3 provides the mean 50% correct recognition points (dB S/N) for the words in each of the eight NAM categories. The data in Figure 3 show that the HLL words required the lowest SNR (−1.3 dB) to obtain 50% correct recognition, whereas the LLH (*L*ow word frequency, *L*ow neighborhood density, and *H*igh neighborhood frequency) words required the highest SNR (2.1 dB) to obtain 50% correct. The range of the mean 50% points from easiest to hardest NAM category was 3.4 dB. The eight NAM categories shown in Figure 3 were subjected to a one-way ANOVA, which revealed a significant effect of NAM category, $F(7, 489) = 6.9$, $p < .001$. Post hoc comparisons using a Bonferroni correction for multiple t tests showed that the easiest words (the HLL words) were not significantly different from the HLH (*H*igh word frequency, *L*ow neighborhood density, and *H*igh neighborhood frequency) or the HHL (*H*igh word frequency, *H*igh neighborhood density, and *L*ow neighborhood frequency) words, which were the next two easiest categories in terms of

Table 3. Number and Percentage of Monosyllabic Words in Each Lexical Category

Lexical Category	Number of Words	Percentage
HLL	63	12.9
HLH	56	11.4
HHL	40	8.2
HHH	85	17.3
LLL	96	19.6
LLH	35	7.1
LHL	44	9.0
LHH	71	14.5
Total	490	100.0

HHH = High word frequency, High neighborhood density, and High neighborhood frequency; HHL = High word frequency, High neighborhood density, and Low neighborhood frequency; HLH = High word frequency, Low neighborhood density, and High neighborhood frequency; HLL = High word frequency, Low neighborhood density, and Low neighborhood frequency; LHH = Low word frequency, High neighborhood density, and High neighborhood frequency; LHL = Low word frequency, High neighborhood density, and Low neighborhood frequency; LLH = Low word frequency, Low neighborhood density, and High neighborhood frequency; LLL = Low word frequency, Low neighborhood density, and Low neighborhood frequency.

recognition performance, but were significantly different from all of the five other categories.

Meanwhile, the two most difficult categories in terms of recognition performance, the LLH and the LHH (*L*ow word frequency, *H*igh neighborhood density, and *H*igh neighborhood frequency) groups, were significantly different only from the two easiest categories, the HLL and the HLH groups. The overall order of the NAM categories from easiest to most difficult is similar to the order of the NAM categories reported by Dirks et al (2001) in that words with high-word frequency were easier to recognize than were the words with low-word frequency.

The only exception is that the LLL (*L*ow word frequency, *L*ow neighborhood density, and *L*ow neighborhood frequency) words in the current study were one-tenth of a decibel easier than the HHH (*H*igh word frequency, *H*igh neighborhood density, and *H*igh neighborhood frequency) words, which is a negligible difference. Dirks et al (2001) also showed similar performance between the LLL and the HHH categories with less than a percentage point difference between the two categories in all three experiments. It is interesting to note that Dirks et al had a 17% difference between the easiest NAM category (HLL, 67.5%) and the hardest NAM category (LHH, 50.4%) when testing listeners with normal hearing in a background of noise. Drawing on an estimated slope of 5%/dB for monosyllabic words, one finds a spread of ~3 dB between the category extremes in the Dirks et al study; in the current study, the range between the easiest NAM category and the hardest NAM category was 3.4 dB.

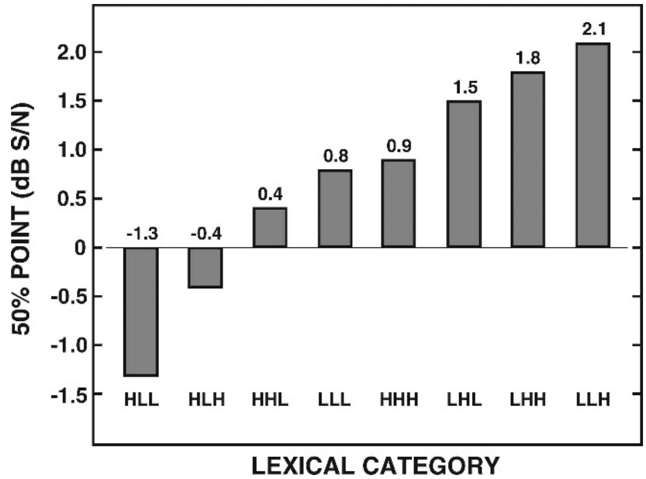

Figure 3. Mean 50% points (ordinate) for each lexical category (abscissa). All lexical categories are named using the same convention so that target word frequency is indicated as low (L) or high (H) in the first position, neighborhood density is indicated as L or H in the second position, and neighborhood frequency is indicated as L or H in the third position.

The data for all 490 monosyllable words listed in the Hoosier Mental Lexicon were analyzed using multiple linear regression to identify predictor variables that accurately predicted the 50% point on the psychometric function. The predictor variables included in the regression analysis were rms; duration; initial phoneme manner, place, and voicing; final phoneme manner, place, and voicing; vowel phoneme; word frequency; word familiarity; neighborhood density; and neighborhood frequency. All categorical variables were dummy-coded[2] for use in the regression analysis. The decibel SNR for the mean 50% point for each of the 490 words was used as the criterion variable.

Blocks of dummy-coded variables were analyzed by using the enter method, which enters all the variables simultaneously. A significant model emerged, $F(28, 461) = 17.41$, $p < .001$, adjusted $R^2 = 0.48$. A model summary with significant predictor variables is listed in Table 4 (left column). Three predictor variables (vowel phoneme, neighborhood density, and neighborhood frequency) did not add significantly to the regression model and are not listed in Table 4. For all of the other predictor variables, the R^2 value, adjusted R^2 value, and the change in R^2 are reported in Table 4. The R^2 value represents the amount of variance in the 50% point that can be accounted for by a specific predictor variable (e.g., rms accounted for 4% of the variance). However, if one takes into account the large number of variables and observations, the adjusted R^2 value is the more appropriate metric to use. Both the R^2 values and the adjusted R^2 values are additive in that the value represents the amount of variance accounted for by the specific variable for that row on the table and all the variables that preceded the variable. Using the adjusted R^2 values, the set of

Table 4. Significant Predictor Variables from the Multiple Regression

Predictor Variable	R^2	Adjusted R^2	R^2 Change
rms Sound Pressure Level	0.04	0.04	0.04
Duration	0.06	0.05	0.02
Initial Phoneme Manner	0.16	0.15	0.10
Initial Phoneme Place	0.22	0.20	0.06
Initial Phoneme Voicing	0.23	0.20	0.01
Final Phoneme Manner	0.37	0.34	0.14
Final Phoneme Place	0.41	0.38	0.04
Final Phoneme Voicing	0.48	0.45	0.07
Familiarity	0.51	0.48	0.03

Note: Significant predictor variables from the multiple regression are listed in column 1. The R^2 values in the second column and the adjusted R^2 values in the third column show the amount of variance accounted for by the addition of each variable so that the total variance accounted for by the model is shown in the last row of the table. The R^2 change values in the fourth column show the variance accounted for by each individual variable. rms = root-mean-square sound pressure.

predictor variables listed accounted for 48% of the variance in the criterion variable (i.e., 50% point).

Interestingly, the majority of the variance was accounted for by the consonant features of the initial and final phoneme, for which the sum of the R^2 change values was 42%. More specifically, the sum for the R^2 change values for the three final phoneme features (manner, place, and voicing), which accounted for a total of 25% of the variance, was greater than for the values when the initial phoneme features were added into the model; the initial phoneme features accounted for only 17% of the variance. Given the differences in variance accounted for by the initial and final phoneme characteristics, examples of words with the same vowel and the same phoneme in either initial or final position were examined to illustrate the greater change in performance as a result of the final phoneme compared to the initial phoneme.

Example word pairs are listed in Table 5; they provide a comparison of performance changes for affricates versus stops in both initial and final phoneme position. This comparison illustrates the greater influence that the final phoneme had on the 50% point. As Table 5 shows, the average difference in 50% points when the initial phoneme was changed from an affricate to a stop was ~3 dB, much lower than the average difference in the 50% points of ~7 dB that was obtained when the final phoneme was changed from an affricate to a stop. This finding—that the final phoneme was more predictive than was the initial phoneme of the whole word 50% point—is consistent with previous reports (e.g., Dubno et al, 1982) that recognition performance was superior in the initial position than it was in the final position. Specifically, if the recognition of the final phoneme were more difficult than recognition of the initial phoneme, then

Table 5. Comparison of the 50% Points for Word Pairs That Differ by Manner for the Initial Phoneme (top panel) and for the Final Phoneme (bottom panel)

Category/Word	Manner	50% point	Difference
Initial Phoneme			
Chair	affricate	−4.9	3.5
Tare	stop	−1.4	
Jail	affricate	1.3	3.0
Pale	stop	4.3	
Average			3.3
Final Phoneme			
Ditch	affricate	−5.8	5.0
Dip	stop	−0.8	
Catch	affricate	−5.3	9.3
Cap	stop	4.0	
Average			7.2

the variance in 50% points as a function of final phoneme features would be greater.

Familiarity was the only lexical variable to account for a significant amount of variance in relation to the 50% point, resulting in a change in R^2 of 3%. The lexical variables of word frequency, neighborhood density, and neighborhood frequency were entered as continuous variables into a regression analysis in addition to being entered as dummy-coded categorical variables (L or H) in a separate regression analysis. Neither regression analysis found that these three lexical variables significantly improved the regression model. This finding was unexpected, given that Dirks et al (2001) identified word frequency as highly predictive of recognition performance for listeners with normal hearing on a speech-in-noise task. One possible reason for the difference between the current study and the Dirks et al study is that the words used in the current study were from the standardized lists commonly used for word-recognition testing (i.e., PB-50, CID W-22, NU–6) and might have been more homogenous in terms of word frequency than the stimulus set used by Dirks et al.

Although the current data set did not replicate the regression results of Dirks et al (2001), the data in Figure 3 do support the broader finding that there are processing advantages for high-frequency words, as evidenced by the words with high word frequency (i.e., HLL, HLH, HHL) being recognized at poorer SNRs (Soloman and Postman, 1952; Rosenzweig and Postman, 1957; Broadbent, 1967).

CONCLUSIONS

The results of this study support the contention that the acoustic properties of an utterance are the foundation for speech perception (Miller and Nicely, 1955). The regression analysis in the current study suggests that the acoustic and phonetic variables associated with bottom-up processing (i.e., rms, duration, and consonant features for the initial and final phoneme) can predict almost half of the variance associated with the recognition performance at the 50% point on the psychometric function for a monosyllabic-words-in-noise test. With the exception of word familiarity, lexical characteristics associated with top-down processing (i.e., word frequency, neighborhood density, and neighborhood frequency) were not found to be significant predictors of word-recognition performance at the 50% point. Because the current results were obtained from listeners with normal hearing, it is possible that the results cannot be generalized to listeners with hearing loss.

The finding that the significant predictor variables in the regression model were almost entirely variables related to bottom-up processing supports the use of monosyllables as a clinical measure for assessing how a listener recognizes speech in noise. Although sentences may be more representative of everyday conversation, a clinician fitting sensory aids seemingly would want the least amount of influence from top-down processes, given that even the best hearing aid fitting will not improve the ability of the individual to use context. Thus, if speech-in-noise testing is used in a pre- and post-hearing-aid-fitting format, then the use of monosyllabic words may be more sensitive to changes in audibility resulting from the hearing aids than contextual-sentence materials would be.

Acknowledgments. This work was supported by the Rehabilitation Research and Development Service, Department of Veterans Affairs, through a Research Career Development award to the first author and a Senior Research Career Scientist award to the second author. The authors would like to thank Jane Smart for the word transcriptions and Judy Dubno and Theresa Chisolm for their comments on the manuscript.

NOTES

1. The 13 lists of words (4 PB-50 lists, 4 CID W-22 lists, 4 NU–6 lists, and the list of nine monosyllabic digits) produced 609 words (12 × 50, plus the nine digits). Of the 609 words, 59 words were in two lists, and 2 words were in three lists, so a net of 546 words was produced.

2. Categorical variables that have more than two levels are dummy-coded with ones and zeros to identify category membership.

REFERENCES

Auer ET, Bernstein LE, Tucker PE. (2000) Is subjective word familiarity a meter of ambient language? A natural experiment on effects of perceptual experience. *Mem Cognit* 28:789–797.

Bell TS, Wilson RH. (2001) Sentence recognition materials based on frequency of word usage and lexical confusability. *J Am Acad Audiol* 12:524–532.

Black JW. (1952) Accompaniments of word intelligibility. *J Speech Hear Disord* 17:409–418.

Bilger RC, Wang MD. (1976) Consonant confusions in patients with sensorineural hearing loss. *J Speech Hear Res* 19:718–740.

Boothroyd A, Nittrouer S. (1988) Mathematical treatment of context effects in phoneme and word recognition. *J Acoust Soc Am* 84:101–114.

Broadbent D. (1967) Word-frequency effect and response bias. *Psychol Rev* 74:1–15.

Cluff MS, Luce PA. (1990) Similarity neighborhoods of spoken two-syllable words: retroactive effects on multiple activation. *J Exp Psychol Hum Percept Perform* 16:551–563.

Cox RM, Alexander GC, Gilmore C. (1987) Development of the Connected Speech Test (CST). *Ear Hear* 8:19S–125S.

Dirks DD, Takayanagi S, Moshfegh P, Noffsinger PD, Fausti SA. (2001) Examination of the neighborhood activation theory in normal and hearing-impaired listeners. *Ear Hear* 2: 1–13.

Dubno JR, Dirks DD, Langhofer LR. (1982) Evaluation of hearing-impaired listeners using a nonsense-syllable test II: syllable recognition and consonant confusion patterns. *J Speech Hear Res* 25:141–148.

Dubno JR, Levitt H. (1981) Predicting consonant confusions from acoustic analysis. *J Acoust Soc Am* 69:249–261.

Egan J. (1948) Articulation testing methods. *Laryngoscope* 58: 955–991.

Etymotic Research. (2005) *BKB-SIN™ Speech-in-Noise Test version 1.03 [CD].* Elk Grove Village, IL: Etymotic Research.

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

Gernsbacher MA. (1984) Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *J Exp Psychol Gen* 113:256–281.

Gordon-Salant S. (1985) Phoneme feature perception in noise by normal-hearing and hearing-impaired subjects. *J Speech Hear Res* 28:87–95.

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

Howes D. (1957) On the relation between the intelligibility and frequency of occurrence of English words. *J Acoust Soc Am* 29: 296–305.

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.

Kirk KI, Pisoni DB, Miyamoto RC. (1997) Effects of stimulus variability on speech perception in listeners with hearing impairment. *J Speech Lang Hear Res* 40:1395–1405.

Kreul E, Bell D, Nixon J. (1969) Factors affecting speech discrimination test difficulty. *J Speech Hear Res* 12:281–287.

Kucera F, Francis W. (1967) *Computational Analysis of Present Day American English.* Providence, RI: Brown University Press.

Luce PA. (1986) A computational analysis of uniqueness points in acoustic word recognition. *Percept Psychophys* 39:155–158.

Luce PA, Pisoni DB. (1998) Recognizing spoken words: the neighborhood activation model. *Ear Hear* 19:1–36.

Luce PA, Pisoni DB, Goldinger SD. (1990) Similarity neighborhoods of spoken words. In: Altmann G, ed. *Cognitive Models of Speech Perception: Psycholinguistic and Computational Perspectives.* Cambridge MA: MIT Press, 122–147.

Mason HM. (1945) *Phonetic characteristics of words as related to their intelligibility in aircraft-type noises.* Washington, DC: Psychological Corp., OSRD Report No. 468.

Miller GA. (1951) *Language and Communication.* New York: McGraw-Hill.

Miller GA, Nicely PE. (1955) An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 27:338–352.

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.

Nusbaum HC, Pisoni DB, Davis CK. (1984) *Sizing Up the Hoosier Mental Lexicon: Measuring the Familiarity of 20,000 Words* (Research on Speech Perception Progress Report No. 10). Bloomington, IN: Indiana University, Department of Psychology, Speech Research Laboratory.

O'Neill JJ. (1957) Recognition of intelligibility test materials in context and isolation. *J Speech Hear Disord* 22:87–90.

Rosenzweig MR, Postman L. (1957) Intelligibility as a function of frequency of usage. *J Exp Psychol* 54:412–422.

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

Soloman RL, Postman L. (1952) Frequency of usage as a determinant of recognition thresholds for words. *J Exp Psychol* 43:195–201.

Sommers M. (1996) The structural organization of the mental lexicon and its contribution to age-related declines in spoken word recognition. *Psychol Aging* 11:333–341.

Sommers MS, Danielson SM. (1999) Inhibitory processes and spoken word recognition in young and older adults: the interaction of lexical competition and semantic context. *Psychol Aging* 14:458–472.

Sommers MS, Kirk KI, Pisoni DB. (1997) Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners I: the effects of response format. *Ear Hear* 18:89–99.

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

Walden BE, Montgomery AA. (1975) Dimensions of consonant perception in normal and hearing-impaired listeners. *J Speech Hear Res* 18:444–455.

Wang MD, Bilger RC. (1973) Consonant confusions in noise: a study of perceptual features. *J Acoust Soc Am* 54:1248–1266.

Wilson RH, McArdle R, Roberts H. (This issue) A comparison of recognition performances in speech-spectrum noise by listeners with normal hearing on PB-50, CID W-22, NU-6, W-1 spondaic words, and monosyllabic digits spoken by the same speaker. *J Am Acad Audiol* 19(6):496–506.

Wilson RH, Morgan DE, Dirks DD. (1973) A proposed SRT procedure and its statistical precedent. *J Speech Hear Disord* 38:184–191.