Published in final edited form as:
Ear Hear. 2010 February ; 31(1): 102–114. doi:10.1097/AUD.0b013e3181b7892f.

Lexical effects on spoken word recognition in children with normal hearing [a]

Vidya Krull[1], **Sangsook Choi**[1], **Karen Iler Kirk**[1,2], **Lindsay Prusick**[1], and **Brian French**[3]

[1]Department of Speech, Language, and Hearing Sciences Purdue University West Lafayette, IN 47907-2038

[2]Department of Otolaryngology-Head and Neck Surgery Indiana University School of Medicine Indianapolis, IN 46202-5230

[3]Department of Educational Leadership and Counseling Psychology Washington State University Pullman, WA 99163-2136

Summary

This paper outlines the development of a theoretically-motivated sentence recognition test for children. Previous sentence tests such as the Lexical Neighborhood Test and the Multisyllabic Lexical Neighborhood Test examined lexical effects on children's recognition of words. In previous studies related to their test development, lexical characteristics were confounded. This study examines independent effects of word frequency and lexical density on a new test of spoken word recognition in children. Results show that word frequency and lexical density influence word recognition both independently, and in combination. Lexical density appears to be more heavily weighted than word frequency in children.

INTRODUCTION

Spoken word recognition tests have been used clinically in the diagnosis and management of individuals with hearing loss for the last 60 years. They are used to assess the effects of hearing loss on communication abilities (Hudgins et al., 1947), to classify the degree and type of hearing loss, to determine candidacy for cochlear implantation (ASHA Technical Report, 2004; Kirk, 2000), to objectively measure the benefit provided by a cochlear implant (CI) or hearing aid, and to provide information that will inform aural (re)habilitation programs (Mackersie, 2002). The goal of such testing is to:

> "…provide a measure of how well listeners understand speech in a controlled environment as a reflection of how they may do in everyday listening situations." (Mendel and Danhauer, 1997 (page 3)).

In research settings, listeners with hearing loss may be administered spoken word recognition tests to evaluate the effectiveness of signal processing strategies, to better understand perceptual processes that support spoken word recognition, or to identify factors that contribute to individual variability. Obviously, no one test can achieve all of these aims.

Traditional clinical tests of spoken word recognition, such as the Phonetically Balanced Kindergarten Word List (PB-K) yield descriptive information concerning speech perception

---

[a] Portions of this research were presented at the 2008 meeting of the American Auditory Society, Scottsdale, AZ, 2008.

Corresponding author: Vidya Krull Department of Speech, Language, and Hearing Sciences Purdue University Heavilon Hall 500 Oval Drive West Lafayette, IN 47907-2038 vganesh@purdue.edu Phone: 765-413-3052 Fax: 765-494-0771.

performance (i.e., the percent of words or phonemes correctly recognized). Such tests may not be suitable for many children with sensory aids because of the severe-to-profound nature of the children's hearing loss, their young age or limited vocabulary. Furthermore, these tests reveal little about the underlying perceptual processes that may contribute to widely varying sensory aid outcomes.

Here we report the first in a series of studies designed to develop a new set of test materials that will help us estimate real-world listening ability and predict benefit from a sensory aid in children with hearing loss. Test development builds upon basic research concerning spoken language processing by listeners with normal or impaired hearing. Milestones in this translational project include selecting or creating new sentence sets, producing equivalent multi-talker test lists and collecting normative data from children with normal hearing or varying degrees of hearing loss. The long-term goal is to provide valid, reliable benchmarks that assist clinicians in making evidence-based decisions about sensory aid use and patient care.

Lexical Effects on Spoken Word Recognition

It is well documented that subject characteristics such as age at onset or degree of hearing loss influence performance on tests of spoken word recognition by children with sensory aids (Seewald et al., 1985; Tillberg et al., 1996). However, the lexical characteristics of the stimulus items also impact performance on such measures. For example, word frequency (i.e., the frequency of occurrence of words in the language) and lexical density (i.e., the number of phonemically similar words to the target) have been shown to affect the accuracy and speed of spoken word recognition performance (Elliott et al., 1983; Luce et al., 1990; Treisman, 1978a; Treisman, 1978b). One measure of lexical similarity is the number of “lexical neighbors” or words that differ by one phoneme from a target word (Greenburg and Jenkins, 1964; Landauer and Streeter, 1973). For example, the words *bat*, *cap*, *cut*, *scat* and *at* are all lexical neighbors of the target word *cat*. Words with many lexical neighbors come from “dense” lexical neighborhoods, whereas those with few lexical neighbors come from “sparse” neighborhoods.

The Neighborhood Activation Model (NAM) of Luce and Pisoni (Luce, 1986; Luce and Pisoni, 1998) offers a two stage account of how the sound pattern of words in memory contributes to the perception of spoken words. According to NAM, a stimulus input activates a set of similar acoustic-phonetic patterns in memory in a multi-dimensional acoustic-phonetic space, with activation levels proportional to the degree of similarity to the target word. This initial activation stage is followed by “lexical selection” among a large number of potential candidates that are consistent with the acoustic-phonetic input. Word frequency is assumed to act as a biasing factor in this model by multiplicatively adjusting the activation levels of the acoustic-phonetic patterns. In lexical selection, the activation levels are then summed and the probability of choosing each pattern is computed based on the overall activation level. Word recognition occurs when a given acoustic-phonetic representation is chosen based on the computed probabilities. Figure 1 illustrates two hypothetical lexical neighborhoods.

Luce and his colleagues (Cluff and Luce, 1990; Luce et al., 1990) showed that word frequency, neighborhood density (i.e., the number of lexical neighbors for the target word), and the neighborhood frequency (i.e., the average word frequency of the words in the lexical neighborhood) all influenced auditory-only spoken word recognition. High-frequency words from sparse neighborhoods were recognized with greater accuracy and speed than low-frequency words from dense neighborhoods. These results demonstrate that the structure and organization of sound patterns in memory affects spoken word recognition by listeners with normal hearing. Similar findings have been demonstrated for adults with hearing loss (Kirk et al., 1997; Sommers, 1996), as well as for children with normal hearing (Charles-Luce and

Luce, 1990; Garlock et al., 2001; Logan, 1992) or with hearing loss (Eisenberg et al., 2002; Kirk et al., 1995; Kirk et al., 1997; Kirk et al., 2000).

Development of the Lexical Neighborhood Tests for Children

Kirk, Pisoni and Osberger (1995) developed new measures that are based on the assumptions of NAM to assess spoken word recognition in children with cochlear implants (CIs). Stimuli on these tests were selected to be familiar to young children with relatively limited vocabularies. To accomplish this, stimuli were drawn from the Child Language Data Exchange System (CHILDES) (MacWhinney and Snow, 1985) which contains transcripts of young children's verbal exchanges with a caregiver or another child. All tokens were drawn from productions by typically-developing children between the ages of 3–5 years, and thus represent early-acquired vocabulary. Based on computational analyses of words in the CHILDES database conducted by Logan (1992), word lists were constructed to allow systematic examination of the effects of word frequency (i.e., number of occurrences of a word within Logan's corpus) lexical density (i.e., the number of words that could be found in the corpus by adding, substituting or deleting one phoneme from the target word) and word length (i.e., monosyllabic vs. multisyllabic words). The following procedures were used to create the word lists. First all proper nouns, possessives, contractions, plurals and inflected forms of words were eliminated from Logan's corpus. This left a total of 994 words in the subset of words produced by 3–5 year old children. Next, median word frequency and lexical density values were computed separately for the remaining monosyllabic and multisyllabic words. Table 1 presents the range and median values of word frequency and lexical density for the subset of words so analyzed. The median values generally served as the cutoff for classifying words as high or low on a given lexical characteristic. In the case of multisyllabic words, the median density value was 0. In that case, words with 0 neighbors were classified as sparse, and words with 1 or more neighbors were classified as dense. Tokens for the tests were then selected so that half of the items were high-frequency sparse words and half were low-frequency dense words. The resulting Lexical Neighborhood Test (Kirk et al., 1995) consists of two lists of 50 monosyllabic words, whereas the Multisyllabic Lexical Neighborhood Test (MLNT) consists of two lists of 24 two-to-three syllable words.

Kirk et al. (1995) used these new measures to examine the effect of lexical characteristics on spoken word recognition performance by children with CIs and to compare their performance on the LNT and MLNT with their performance on the PBK. Recognition accuracy was significantly greater for high-frequency sparse words than for low-frequency dense words. Furthermore, performance was consistently higher on the lexically-controlled word lists than on the PBK. These early results demonstrated that pediatric CI users are sensitive to the acoustic-phonetic similarity among words, that they organize words into similarity neighborhoods in long-term memory, and that they use this structural information in recognizing isolated words in an open-set response format. The results also suggested that the PBK underestimated the participants' spoken word recognition abilities, perhaps because of the vocabulary constraints inherent in creating phonetically balanced word lists.

Kirk, Sehgal and Hay-McCutcheon (2000) compared pediatric cochlear implant recipients' familiarity with words on the PBK, LNT, and MLNT using parent ratings on a seven-point scale. On this scale, “1” represents extreme unfamiliarity with a word and “7” represents extreme familiarity. Results showed a significant difference in familiarity between words on the three lists. Table 2 presents the mean familiarity scores for the high-frequency sparse and low-frequency dense words on the LNT and MLNT as well as on the PBK. Words on the LNT were rated as most familiar followed by the MLNT and the PBK, respectively. Poor performance on the PBK may result, in part, because children with profound deafness are less familiar with the test item vocabulary. There were no significant differences in familiarity between the high-frequency sparse and the low-frequency dense words on the LNT and MLNT. Additionally, word familiarity was significantly related to chronological age for the MLNT and PBK, but not the LNT. These results provide further support for the appropriateness of the LNT as a spoken word recognition test for children of widely varying ages.

Eisenberg and colleagues (2002) created lexically-controlled sentences using monosyllabic or bisyllabic words drawn from the same subset of the CHILDES database used to develop the LNT and MLNT. However, they allowed the use of proper names. The lexical statistics for the combined group of monosyllabic and bisyllabic words are presented in Table 1. As classified by Eisenberg et al., high frequency words had frequency values >2 in Logan's corpus whereas values of ≤ 2 were low frequency. Sparse words had density values ≤2 in Logan's corpus whereas dense words had density values >2. Three key words were used in constructing each of the 5- to 7-word sentences. These sentences were combined into two lists, each with five practice and 20 test sentences that were syntactically correct but semantically neutral. One list contained sentences with high-frequency sparse key words; the other list contained low-frequency dense key words.

The lexically-controlled sentences were audio-recorded by the same female talker and used in three experiments. In Experiment 1, 48 children with normal hearing between the ages of 5–12 years repeated the isolated words and sentences at one of six different levels to generate performance-intensity functions. In Experiment 2, 12 normal hearing children aged 5–14 years repeated the words and sentences under spectrally degraded conditions intended to simulate CI speech processing. In Experiment 3, 12 children with CIs aged 5–14 years repeated the unprocessed stimuli. Children also completed a test of vocabulary recognition. In all three experiments, sentences containing high-frequency sparse key words were recognized better than sentences containing low-frequency dense key words. Sentence scores were significantly higher than word scores for the children with normal hearing and nine high-performing children with CIs. Three low-performing children with CIs showed the opposite pattern for isolated word and sentence stimuli. A statistically significant relationship was observed between chronological age and sentence scores for children with normal hearing who heard degraded speech. For children with CIs, the relationship between language abilities and spoken word recognition was strong and significant. This result demonstrates the influence of linguistic knowledge on phonological processing of words.

Independent Effects of Word Frequency and Lexical Density

In the theoretically-motivated word lists for children described above, the factors of word frequency and lexical density were confounded. In contrast, other authors have used the assumptions of NAM to create adult word lists (Dirks et al., 2001) or sentence lists (Bell and Wilson, 2001) in which the two factors were orthogonally combined. For example, Bell and Wilson created 320 sentences, each with three key words drawn from one of four lexical categories: high-frequency sparse, high-frequency dense, low-frequency sparse, and low-frequency dense. They administered these sentences in quiet and in noise to adults with normal hearing. Bell and Wilson found significant independent effects of word frequency and lexical density, as well as an interaction between the two factors.

As part of our test development, we followed the model of Bell and Wilson by creating a pediatric sentence set in which word frequency and lexical density were orthogonally combined. The purpose of the present study was to examine the independent contributions of word frequency and lexical density to the perception of key words in this new sentence set.

MATERIALS AND METHODS

Participants

Thirty four children between the ages of 5 and 12 years participated in this study. This age range was chosen because it matches the age range used by Eisenberg et al. (2002). Ten of the children participated in a pilot study and 24 participated in the main study (Table 3). Participants did not overlap between the two studies. The median age was 7 years, 5 months for children in the pilot group and 9 years, 1 month for those in the main study. All children had normal hearing in both ears (screened at 20 dB HL at octave frequencies from .25 to 8 kHz), no known history of speech and language difficulties, and were native speakers of American English.

The PPVT-III Form B (Peabody Picture Vocabulary Test, 3rd edition, Dunn and Dunn, 1997) was administered to ensure that all children had normal, age-appropriate receptive vocabulary. If any articulation errors were noted during interactions with the child, the Goldman Fristoe Test of Articulation (GFTA-2, Goldman and Fristoe, 2000) was administered to ensure that articulation skills were age-appropriate. Age-appropriate articulation errors were noted for interpretation of verbal responses. Informed consent was obtained from the parents of the participants and informed assent was obtained from children aged 7 years or older per the protocol approved by the Institutional Review Board at Purdue University. All participants were recruited from the West Lafayette/Lafayette (Indiana) area and the Purdue University community and were paid for their participation.

Stimulus Generation

In our study, we used the 50 high-frequency sparse and low-frequency dense sentences (25 each) created by Eisenberg et al. (2002). We created 50 additional lexically-controlled sentences (25 low-frequency sparse and 25 high-frequency dense) from the remaining words in the Logan (1992) corpus used by Kirk et al. (1995). All the words are chosen from the same (original) dataset of words derived from the CHILDES database, based on words produced by normal-hearing children ages 3–5 yrs. We used the same word frequency and lexical density values as Eisenberg et al. (2002) to classify words as high- vs. low-frequency or sparse vs. dense. Table 4 presents the mean lexical values for the words we used; these values are taken from the Logan corpus. The resulting 100 sentences were syntactically correct but semantically neutral (and hence expectedly low in word predictability). Within a sentence, all key words belonged to the same lexical category. A list of the 300 key words used to create the sentences can be found in the Appendix.

To ensure that the sentences were low in predictability, we obtained measures of predictability from staff members in the Department of Speech, Language, and Hearing Sciences at Purdue University. The individuals were divided into three groups. Each group was given a list of sentences with one of the key words missing in each sentence and asked to fill in the blank with an appropriate word. For example, the first group was presented with a list of sentences with the first key word missing; the second group was presented with a list of sentences with the second key word missing, and so on. If two or more individuals in a group correctly guessed the missing key word in a sentence, the sentence was revised and the process was repeated.

A speech feature analysis was carried out on the 300 key words (from 100 sentences). Figure 2 illustrates that the distribution of speech features is similar across the four lexical categories.

Recording of Stimuli

The lexically-controlled sentences and isolated key words were digitally recorded by a female native speaker of American English in a double-walled IAC sound-attenuating booth (ANSI

S3.1-1991 (R1999)). The recordings were made using a Crown headworn microphone (CM-312A/E) at a sampling rate of 44.1 kHz, with a 16-bit mono recording. The microphone was placed approximately 8 inches away from the speaker's mouth. The speaker repeated each word and sentence several times. Recordings with undesirable sound quality (e.g. clicks, vocal fry) were rejected and the best example of each recorded word and sentence was selected for use in the study. Stimuli were re-sampled to yield 16-bit, 22 kHz, mono-recordings using Adobe Audition®. The RMS amplitude of each word and sentence was equalized to 65 dB SPL using the Level16 program (V2.0.3.) developed by Tice and Carrell (1998).

Pilot Study: Determining Signal-to-Noise Ratio (SNR) for Participant Testing

The purpose of the pilot study was to select four SNRs such that they would optimally span the performance intensity functions for all children in the age range of interest. Because Eisenberg et al. found a significant correlation between chronological age and sentence scores, children in the pilot study were separated into two groups: Group 1 (5–7 year olds) and Group 2 (8–12 year olds).

Stimuli and Instrumentation—Testing was performed in noise in order to eliminate the possibility of ceiling effects when testing children with normal hearing in quiet. For efficient masking, a spectrally-matched noise masker was created by matching the spectrum of the noise to the long term average spectrum of the stimuli. The procedure for creating the noise was as follows: 1) all sentence stimulus files were concatenated; 2) the average spectrum level for the concatenated sentences was analyzed using a Fast Fourier Transform (FFT) (1024-points with Blackmann Harris window); 3) frequency and amplitude values obtained from the spectrum analysis were used to create a FFT filter mimicking the shape of the average spectrum for the sentences; 4) white noise was processed through the FFT filter. The spectrally-matched noise for the word stimuli was created separately using the same procedure. The duration of the noise was based on the duration of the longest stimulus; this was calculated separately for words and sentences. The noise started 100 ms prior to the onset of individual speech stimuli and ended 100 ms after the offset of the longest speech stimulus.

Procedure—The Behavioral Auditory Research Tests (BART) software (Boystown National Research Hospital) was used to implement an adaptive testing procedure (Levitt, 1971). Individual stimulus files were mixed digitally with noise and then routed though the 24-bit sound card (Audiophile 2496, M-Audio) to an amplifier (NAD Stereo Amplifier 3020A). Stimuli were presented binaurally through SONY MDR-D777 headphones. During testing, speech was held constant at 65 dB SPL and the noise was adaptively varied to converge on the 30% and 70% words correct performance level. These levels were chosen to span the performance-intensity function for the new test materials. Each participant completed the procedure four times in order to calculate SNR separately as a function of stimulus context (isolated words vs. key words in sentences) within each convergence level (30% and 70%). Different sentence and words lists were used for each run. The lexical categories were counterbalanced throughout each list to try to insure that the participants heard a similar number of tokens from each category.

Participants were seated in front of a monitor inside a double walled IAC sound-attenuating booth. An examiner sat in the booth with the child to provide encouragement, if necessary. Participants were instructed to listen to speech through headphones and repeat the words or sentences into a microphone. They were encouraged to guess if they were unsure of the response. They were alerted to each stimulus via the computer monitor, which flashed the word `ready' prior to stimulus presentation. An examiner seated outside the booth scored the verbal responses online. For sentence stimuli, only the key words within the sentence were scored.

Responses were used as inputs to the adaptive program in BART. As the signal level was fixed, a SNR corresponding to the desired performance was calculated from the threshold for noise.

Results—Table 5 shows the average SNR corresponding to the 30% and 70% performance criteria as a function of age group and stimulus context. A higher SNR yielded better word recognition performance for both words and sentences regardless of age group. When presented with sentences, children in the younger age group required a slightly greater SNR, or better audibility of the signal, to achieve the same criterion as children in the older age group. As shown in Table 5, this difference was the greatest for sentences at the 70% convergence level. However, when they were presented with words, the SNR needed to achieve similar performance was comparable across age groups. Based on this information, SNRs of −2, 0, 2, and 4 dB were chosen to make sure that performance for either age group would not be affected by floor or ceiling effects.

Main Study

Stimuli and Instrumentation—Stimuli and instrumentation in the main experiment were similar to that in the pilot study, with the exception that for an individual subject the SNR was fixed at one of the four SNRs determined by the pilot study. In addition, a Shure PG185 lavalier microphone was used to record spoken responses to a digital audio recorder (Marantz PMD-671). The recording was subsequently used to verify the scoring accuracy and to assess inter-rater agreement.

Procedure—Participants were divided into four groups (n=6 per group), corresponding to one of four SNRs (−2, 0, 2 and 4 dB), with age balanced across groups (Table 3). In the main study, each participant heard all 100 sentences and the 300 key words. Presentation of sentence and word lists was counterbalanced across participants within each group. Each participant listened to the isolated words and sentences at one SNR. The protocol was similar to the pilot study with the exception that two examiners (one inside, and one outside the booth) independently transcribed the participant's response in order to calculate inter-rater agreement.

Scoring and Agreement—Initial point-by-point agreement ratings and verification of scoring by listening to the recordings revealed that when discrepancies occurred, the examiner inside the booth was invariably correct. Hence the scoring by the examiner seated inside the booth was taken to be final in subsequent cases. To assess inter-rater agreement, an analysis was conducted. A naïve research assistant with no prior involvement in the data collection listened to recordings from 5 randomly selected participants (about 20% the sample) and scored them. These scores were then compared to the original scores from the examiner seated inside the booth. Agreement was calculated for each participant as follows:

Inter-rater Agreement = Number of agreements / Total number of items

The analysis revealed that inter-rater agreement remained above 98% in all the samples that were checked.

Data Analyses—Multivariate analyses of variance (MANOVA) were performed to examine the effects of SNR, word frequency, and lexical density on spoken word recognition performance. SNR (4 levels) was included as a between-subject factor; word frequency (high vs. low) and lexical density (sparse vs. dense) were included as within-subject factors. MANOVAs were conducted separately for word and sentence scores. Effect sizes (η^2) are reported to assist with interpretation of the statistical test results. Effect sizes around 0.01 represent small effects, 0.06 represent medium effects, and 0.14 represent large effects (Cohen, 1988). Additionally, correlational analyses were carried out to examine the relationship between spoken word recognition performance and computational values for key word frequency and lexical density.

RESULTS

Standard data screening and assumption checking was conducted prior to the analysis for the word and sentence data. Levene's test for homogeneity of variance was not significant for any of the dependent variables (ps > 0.01). The assumption of homogeneity of covariance matrices, tested via Box's M, were not significantly different ($p > 0.01$) for both words and sentences. We concluded multivariate normality could be assumed with (a) the normalized estimate for multivariate kurtosis values being −1.06 and 0.63 for words and sentences respectively, and (b) the proportion of Mahalanobis Distances not deviating sufficiently from the 50th percentile of the chi-square distribution for words (0.54) and sentences (.39). Meeting these assumptions, we proceeded with the analyses.

Noise Effects

SNR was found to have a significant effect on the recognition of both isolated words (F (1, 20) =37.946, $p < 0.01$, partial $\eta^2=0.851$) and key words in sentences (F (1, 20) =38.601, $p < 0.01$, partial $\eta^2=0.853$). Average group scores increased with increases in SNR used for test administration (Figures 3 and 4). To evaluate if this improvement was similar for all step-wise increases in SNR, a post-hoc analysis was performed. Tukey's Honestly Significant Difference (HSD) procedure was used to follow-up significant overall differences. This is a conservative procedure for controlling Type I error rate (Field, 2005) and is probably the most often used procedure when making multiple pairwise comparisons. A Tukey test revealed significant differences (<0.05) between all SNRs except 2 and 4 dB for sentences and words. Significant mean differences ranged from 7.67% between 0 and 2 dB SNR to 22.63% for −2 to 4 dB SNR for words. Significant mean differences ranged from 10.79% between 0 and 2 dB SNR to 27.29% for −2 to 4 dB SNR for sentences. Performance improved up to 2 dB SNR, after which further increase in SNR did not change performance significantly.

Word Frequency Effects

There also was a significant main effect of word frequency on the recognition of key words in isolation (F (1, 20) =28.2, $p<0.05$, partial $\eta^2 = 0.585$) and in sentences (F (1, 20) =8.727, $p<0.05$, partial $\eta^2 =0.304$). When words were presented in isolation, performance for high-frequency words was significantly better than that for low-frequency words at all SNRs, as shown in Table 6. However, in a sentence context, although a significant effect of word frequency was noted on children's performance (Table 7), the result was a bit more complicated due to an interaction between word frequency and lexical density as discussed below.

Lexical Density Effects

Finally, there was a main effect of lexical density. That is, words from sparse neighborhoods were recognized with significantly greater accuracy than words from dense neighborhoods. The mean scores for the sparse conditions were higher than those for the dense conditions, both for isolated words (Table 6) and sentences (Table 7), irrespective of SNR. The effect of lexical density was significant for both words (F (1, 20) =140.857, $p<0.05$, partial $\eta^2 = 0.876$) and sentences (F (1, 20) =59.732, $p<0.05$, partial $\eta^2 = 0.749$).

Lexical Density * Noise

A significant interaction was found between lexical density and SNR for words presented in isolation (F (3, 20) =3.911, $p<0.05$, partial $\eta^2 = 0.370$). Increasing the SNR by 2 dB yielded greater improvements in word recognition at less favorable than at the more favorable SNRs (see Figure 3). Following recommendations of Field (2005), the interaction was dissected with simple effects analyses which controlled for Type I error. Results revealed that the sparse word scores differed significantly ($ps < 0.01$) between SNRs of −2 to 2 dB, of −2 to 4 dB, and of 0 to 4 dB. Dense word scores were significantly different ($ps < 0.01$) for the comparisons of SNRs −2 to 2 dB, −2 to 4 dB, 0 to 2 dB and 0 to 4 dB. In all cases, the high SNR condition resulted in higher word scores. No corresponding interaction was evident for sentences.

Word Frequency * Lexical Density

The MANOVA revealed a significant interaction between word frequency and lexical density, for both words presented in isolation ($F\ (1, 20) = 4.573$, $p<0.05$, partial $\eta^2 = 0.186$) and in a sentence context ($F\ (1, 20) = 81.33$, $p<0.01$, partial $\eta^2 = 0.803$). However, effect size and power was larger for sentences compared to words. Dissecting the interaction with simple effect analyses (Field, 2005) revealed that recognition accuracy differed significantly ($p< 0.01$) for all combinations of interactions between word frequency and lexical density (high-frequency sparse, high-frequency dense, low-frequency sparse, and low-frequency dense) for both words and sentences. In addition, sparse words yielded greater accuracy than dense words, irrespective of word frequency, both in isolation and in a sentence context (Tables 6 and 7).

Correlational Analyses

Correlational analyses were conducted to examine the relationship between word recognition and the lexical characteristics of the key words. As shown in Table 8, performance scores were not significantly correlated with word frequency values. There was a modest, but significant negative correlation between lexical density and the recognition of words in isolation and in sentences. That is, as the number of lexical neighbors to a target increased, recognition accuracy decreased.

DISCUSSION

This study examined whether the lexical factors of word frequency and density independently influenced spoken word recognition performance on a new, theoretically-motivated sentence recognition test for children. We also investigated the influence of context (sentence vs. words) on the recognition of lexically-controlled key words. Children were tested in spectrally-shaped noise at SNRs chosen to span the performance-intensity function for these test materials.

Our findings offer further evidence that supports the Neighborhood Activation Model (Luce and Pisoni, 1998). Specifically, word frequency and lexical density both significantly influenced the auditory-only recognition of key words in this group of children with normal hearing. Word frequency effects differed for isolated words and sentences. When presented with the key words in isolation, high frequency words were recognized with greater accuracy than low frequency words, all other things being equal. The same was not true in the sentence context. A more complex pattern was evident in the interaction effects, as discussed below.

In contrast, the effect of lexical density was consistent regardless of context. That is, key word recognition was significantly better for words that came from sparse neighborhoods compared to those from dense neighborhoods both in isolated words and in sentences. Also this result is consistent with the predictions of NAM, i.e. when the neighborhood is dense, the probability of correct target word identification decreases. Listeners need to make use of fine acoustic-phonetic cues in order to be able to identify the target word from similar sounding neighbors. Initially, children may be able to recognize words on the basis of relatively gross spectral cues because of their limited vocabulary. However, as words are added to the lexicon, the recognition task becomes more difficult, especially for target words from dense neighborhoods.

The larger the neighborhood, the finer the acoustic-phonetic distinctions that are required to correctly select the target word.

The word frequency and lexical density effects demonstrated in the current study are in agreement with those previously reported for children (Eisenberg et al., 2002; Kirk et al., 1995). However, word frequency and lexical density were confounded in these earlier investigations. That is, all words were either high-frequency sparse or low-frequency dense. In the current study, we followed the model of Bell and Wilson (2001) who orthogonally combined the two factors when creating sentence lists for adults. This allowed us to more carefully examine the relative contribution of the two factors to the children's perception of spoken words. Our results will be discussed in relation to those of Bell and Wilson (2001).

As in the present study, Bell and Wilson found that both main effects were significant; performance was significantly better for high frequency than low frequency words, and for sparse rather than dense words. Similarly, both studies revealed an interaction between word frequency and lexical density effects. However, the pattern of results across all four categories differs between the two investigations. Bell and Wilson's data revealed that word recognition performance was best for the high-frequency sparse category, followed respectively by the categories of high-frequency dense, low-frequency sparse, and low-frequency dense. Thus, their results suggested that recognition performance in adults was influenced more by word frequency than lexical density.

In the current study we see a somewhat different ranking when children were asked to identify the key words in isolation. The best performance was for the high-frequency sparse category, followed respectively by the categories of low-frequency sparse, high-frequency dense, and low frequency dense. In other words, recognition for single key words from sparse neighborhoods was better than that for single key words from dense neighborhoods, regardless of word frequency. The ranking of the lexical categories differed when we looked at key word recognition in a sentence context. In that context, key word recognition was best in the high-frequency sparse category, followed respectively by the categories of low-frequency sparse, low-frequency dense, and high-frequency dense. In other words, when key words in the sentences came from sparse neighborhoods, the children's performance was significantly better for high frequency than low frequency words. However, when the key words in sentences came from dense neighborhoods, they recognized low frequency words with greater accuracy than high frequency words. For both words and sentences, these results suggest that recognition performance in children is weighted more by lexical density than word frequency.

Why do our results differ as a function of context? And why do our sentence data differ from those of Bell and Wilson? There are several possible reasons. It is possible that in creating syntactically correct and semantically neutral sentences, we may have created some sentences that were more or less predictable than others. To determine whether or not this was the case, we examined the performance for individual sentences at different SNRs. We found that some sentences were often recognized correctly (mean key word correct ≥ 2.75) whereas others were very not (mean key word correct was <1) across all SNRs. It is possible that scoring these sentences could have skewed the results. However, removal of such sentences from the analysis did not change the results in any significant manner. Furthermore, care was taken in sentence development to revise any sentence that was highly predictable.

The difference in category rankings between the two studies may be due, in part, to the different sources of material used to create the sentence lists. In the Bell and Wilson study, words were drawn from a database created in the Speech Research Laboratory at Indiana University, Bloomington (Nusbaum et al., 1984). This database contains all 20,000 words in a computer-readable version of Webster's Pocket Dictionary (Pisoni et al., 1985) and density data are based on an analysis of this corpus. Word frequency values in this database refer to the frequency of occurrence per one million words of text (Kucera and Francis, 1967). In contrast, words for the pediatric sentence test were drawn from productions of young children. Density and word frequency values were computed from a much smaller set of stimuli and the corresponding range of word frequency and density values are greatly reduced compared to the adult corpus. It is possible that in creating our sentences, we did not use a frequency range broad enough to reveal the effects of word frequency in children with normal hearing. This may be why density effects were more heavily weighted in the current investigation. Furthermore, the restricted frequency range also may explain the lack of a significant correlation between word frequency and spoken word recognition scores.

It is also possible that combining the key words in a sentence context had a multiplicative effect on the way in which lexical factors influenced word recognition. In a sentence format, the combined effect of word frequency and lexical density attributed to the three key words may not be a simple additive effect. The level of difficulty of the spoken word recognition task may increase as a result. Further, the neighborhood frequency could also have played a role. Logan (1992) reported that high frequency words typically have a greater number of neighbors (dense neighborhood) than low frequency words. Further, the neighbors of high frequency words tend also to be high frequency. Assessments of the effects of neighborhood frequency may be informative in future studies.

As pointed out by one of our reviewers, it is possible that word frequency estimates computed from the productions of very young children are not valid for the age range of the children in the current investigation (5–12 years). There is some evidence that children differ from adults in the manner in which they organize and access lexical information, and that lexical effects vary by age (Charles-Luce and Luce, 1990; Garlock et al., 2001). Charles-Luce and Luce (1990) suggested that younger children with limited vocabularies may use broad phonetic categories (based on manner and place of articulation) to recognize words; older children who have larger vocabularies must be able to make fine acoustic-phonetic details in the speech signal in order to discriminate target words from their competitors. It is likely that frequency and density values for the key words in our sentences would be different if a corpus of words produced by children aged 5–12 years had been analyzed. Alternatively, we could have chosen participants whose age more closely matched the ages of the children who produced the words in the first place. Because we incorporated the sentences created by Eisenberg et al. (2002) into our larger sentence test, we wanted to test participants of the same age as in their earlier study. Future investigations are needed to examine lexical effects on sentence recognition in children younger than 5 years. The performance of older children with hearing loss also should be examined. Because children with hearing loss often have delayed language skills, their results should be compared to age-matched, as well as language-matched peers.

Not surprisingly, the current results showed that performance improved as the SNR improved. Lexical effects on spoken word recognition were largest at poorer SNRs, suggesting that lexical effects may be most pronounced under degraded listening conditions, such as in the presence of background noise, or when children with hearing loss listen to speech through a sensory aid. Lexical effects may be more evident at poorer SNRs because the degraded auditory signal limits access to the fine spectral details of speech. However, as pointed out by a reviewer, lexical effects may come into play relatively late during processing of spoken words. Thus, degraded listening conditions may result in slowed perceptual processing, giving time for lexical effects to emerge. The inclusion of reaction time data in future studies should provide further information concerning lexical effects on sentence recognition. Reaction time data may reveal differences among lexical categories with similar degrees of recognition accuracy.

Results from our pilot data show that to achieve criterion performance (30% or 70%), a slightly greater SNR was required for sentences than for isolated words in three out of four comparisons (Table 5). These results would appear to be contrary to those of Miller and colleagues (1951) who reported that recognition accuracy at a given SNR was greater for words in sentences than for isolated words. However, given the nature of the adaptive task we used, the effects of context cannot be determined. The adaptive program terminated when children achieved a preset criterion performance (30% or 70%). As a result, for the same criterion performance, a child may have heard a different number of stimuli for sentences and words. Further, because the lists were randomized, the keys words in the sentence and word lists were likely to be different, so the effect of context cannot be examined. Secondly, Miller et al. (1951) explain that the advantage of context is mainly due to the `degree of restriction' or the predictability of key words in sentences. In our study, due care was taken to maintain their unpredictability; for example a sample sentence was: “<u>A monster</u> was <u>cleaning</u> the <u>refrigerator</u>.”

The present results also have clinical implications. Current tests of speech perception like the LNT and MLNT use controlled lexical properties to tap the perceptual processes that underlie spoken word recognition. However, these tests use stimuli in which word frequency and lexical density covary. It is therefore difficult to parse the effect that each of these factors separately has on spoken word recognition. The use of orthogonal, lexically-controlled sentences allows us to assess the effect of each of these factors on receptive communication. Furthermore, sentence test material may better represent everyday speech than isolated words.

We currently are developing sentence test materials that will assess the perceptual processes governing speech recognition in children in greater detail. We have recorded 10 talkers producing the current 100- sentence list in an audiovisual format, along with Bell and Wilson's larger sentence set for adults. Studies are underway to determine the intelligibility of each talker in each presentation format (auditory-only, visual-only and auditory-visual). These data will be used to create equivalent multitalker lists within each presentation modality. As part of test development, we also plan to collect normative data from adults and children as a function of hearing status (normal vs. varying degrees of hearing loss) and type of sensory aid (hearing aid or CI) for listeners with hearing loss. These new norm-referenced, multimodal sentence tests should provide important insights into differences in spoken word recognition and the enormous variability noted among individuals with hearing loss. In addition, they should lead to better diagnosis, evaluation and assessment paradigms for listeners with hearing loss. Information obtained from these new measures should prove useful in selecting sensory aids and in developing intervention programs that are targeted to an individual's specific needs. Finally, these tests should better reflect real-world listening abilities than traditional speech discrimination tests employing highly constrained auditory stimulus materials.

Acknowledgments

This work was supported in part by grant numbers R01 DC008875 and R01 DC00064 from the National Institute on Deafness and Other Communication Disorders. The content is solely the responsibility of the authors and does not necessarily reflect on the official views of the National Institute on Deafness and Other Communication Disorders. We wish to thank Dyanna Hamstra, Julie Bassetto, and Rachel Burnett for assistance in data collection and scoring. We would also like to thank Jennifer Karpicke, Kyle Krull, the section editor and the three anonymous reviewers for valuable comments on an earlier version of the manuscript.

This research was supported by NIH grants R01 DC008875-01 and RO1 DC000064-18.

APPENDIX

TABLE A1

High-frequency sparse words

Word	Word frequency	Lexical density	Word	Word frequency	Lexical density	Word	Word frequency	Lexical density
daddy's	4	1	finger	11	0	door	3	1
mouth	5	0	school	7	1	seven	5	1
turning	5	0	friend	8	1	eggs	3	1
sky	5	1	thinks	4	1	street	3	1
doesn't	4	0	lipstick	3	0	just	6	0
tonight	3	0	give	7	1	grey	11	1
crazy	5	0	monkey	7	0	shoelace	3	0
turtle	11	1	juice	4	1	wonder	4	1
away	7	1	wash	3	1	brought	3	1
Nana	3	0	ducks	8	1	food	9	0
getting	4	1	myself	6	0	which	3	0
glasses	10	0	draw	3	1	space	4	0
didn't	6	0	little	55	1	black	9	1
scratch	4	1	snake	9	1	It's	57	2
orange	9	0	open	8	1	watch	54	1
kind	4	1	green	10	0	fish	8	0
airplane	4	0	first	3	0	let's	6	1
brown	11	1	string	9	0	gas	3	1
stand	4	0	stay	3	1	from	8	0
broken	6	0	pocket	3	0	girl	8	1
truck	4	1	please	3	1	takes	4	1
children	12	0	help	3	1	milk	3	1
cried	3	1	puzzle	4	0	break	9	1
farm	15	1	don't	51	1	when	11	1
broken	9	1	scribble	3	1	jump	5	1

TABLE A2

High-frequency dense words

Word	Word frequency	Lexical density	Word	Word frequency	Lexical density	Word	Word frequency	Lexical density
penny	5	10	root	3	2	this	74	2
book	3	6	missed	6	2	make	16	7
old	38	3	duck	19	5	pig	14	4
catch	4	3	red	13	8	hold	5	4
egg	4	4	swing	4	2	won't	11	3
tie	4	14	lock	8	6	pull	10	8
read	4	5	cat	4	13	oar	8	10
sing	4	8	shoe	5	8	dog's	3	3
mother	4	2	baby	21	3	knife	3	3
cake	3	8	cut	3	4	wet	4	8
still	3	3	star	4	3	think	12	6
smells	3	2	hear	34	10	daddy	8	3
band	7	6	cow	6	4	sew	10	10
goes	19	7	rap	10	3	why	22	2
eat	8	8	made	4	2	sit	5	4
push	8	2	boats	4	2	home	6	4
ice	5	6	white	11	2	boy	16	4
dock	8	5	may	10	14	caught	13	11
Paul	17	11	throw	3	3	eel	46	19
hot	3	11	something	18	2	should	4	2
pie	4	11	fire	3	5	love	4	4
threw	5	2	cold	5	3	colors	8	2
button	4	4	six	7	2	tea	3	14
Kent	3	3	shake	3	5	bright	5	3
Here's	9	5	chicken	6	3	now	85	6

TABLE A3

Low-frequency sparse words

Word	Word frequency	Lexical density	Word	Word frequency	Lexical density	Word	Word frequency	Lexical density
visitors	1	0	bubble	1	0	popsicle	1	0
stretched	1	0	pretty	1	0	pants	1	1
dinner	1	0	Jacob	1	0	really	1	1
Saturday	1	0	blocks	1	1	soft	1	0
finish	2	1	outside	1	0	teddy-bear	1	0
costume	2	0	catching	1	0	waiting	1	0
camp	2	1	bugs	1	1	closet	1	0
needs	2	1	dirty	2	0	warmed	1	0
barn	1	1	lion	1	1	cousins	1	0
family	1	0	tricked	2	1	mirror	1	0
Germany	1	0	alligator	2	0	Peter	1	0
pictures	2	0	chief	2	0	wheel	1	0
glove	1	1	cartwheel	2	0	summer	2	1
carry	1	1	Joanne	2	0	haven't	1	1
pickle	1	1	Adam	1	0	scribbled	2	1
business	1	0	banjo	1	0	downstairs	1	0
strange	1	0	garage	1	0	cranky	2	0
commercial	2	0	monster	1	0	beetle	1	0
forgot	1	0	cleaning	1	0	wanted	1	0
bicycle	2	0	refrigerator	1	0	fixed	2	1
scratched	2	1	gramma	1	1	balloon	1	0
uses	1	0	painting	1	0	pencil	2	0
pepper	1	0	nursery	1	0	easy	1	0
salad	1	0	steam,	1	0	point	1	0
shape	1	1	freezing	1	0	river	1	0

TABLE A4

Low-frequency dense words

Word	Word frequency	Lexical density	Word	Word frequency	Lexical density	Word	Word frequency	Lexical density
Billy	1	2	played	1	4	books	2	3
walked	1	2	chickens	1	2	gum	1	7
wrong	2	4	ever	1	2	tiny	1	3
Marie	1	2	find	1	3	box	2	2
thought	1	6	toys	1	3	tummy	1	2
were	1	9	Grampa	1	2	ten	2	5
save	2	2	laughed	1	2	days	1	3
teeny	2	3	goats	1	2	start	1	2
bowl	1	6	dad	2	8	walking	1	2
pick	1	5	came	1	6	seat	1	6
books	2	3	hello	2	2	taught	1	7
Eddie	1	3	boys	1	3	us	2	5
Tommy	1	2	turns	1	2	trick	2	2
once	1	4	locking	1	2	worm	1	3
hat	2	13	many	2	6	stuck	1	3
tell	2	5	kids	1	2	pool	1	7
sleep	1	2	learn	2	2	guess	2	3
belly	1	3	lost	2	4	were	1	9
bunny	1	3	mommy's	1	2	rain	1	2
hid	1	9	ring	2	6	cups	1	2
room	1	2	knows	2	6	pink	2	5
likes	2	2	leave	2	2	bag	1	5
share	1	10	money	2	7	both	2	4
butter	1	4	piggy	1	2	cats	2	3
son	1	6	moved	1	2	mine	2	10

REFERENCES

American National Standards Institute. Maximum permissible ambient noise levels for audiometric test rooms. ANSI; New York: 1999. (ANSI S3.1-1991 [R1999])

American Speech-Language-Hearing Association. Cochlear Implants. 2004 [Technical Report]. Available from www.asha.org/policy.

BART (Behavioral Auditory Research Tests). Boystown National Research Hospital Computing Core. 2006. [Computer software](supported by NIH grant P30-DC004662)

Bell TS, Wilson RH. Sentence recognition materials based on frequency of word use and lexical confusability. Journal of American Audiology 2001;12:514–522.

Charles-Luce J, Luce PA. Similarity neighborhoods of words in young children's lexicons. Journal of Child Language 1990;17:205–215. [PubMed: 2312642]

Cluff MS, Luce PA. Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation. Journal of Experimental Psychology: Human Perception and Performance 1990;16:551–563. [PubMed: 2144570]

Cohen, J. Statistical power analysis for the behavioral sciences. Erlbaum; Hillsdale, NJ: 1988.

Dirks DD, Takayanagi S, Moshfegh A, Noffsinger D, Fausti SA. Examination of the neighborhood activation theory in normal hearing and hearing-impaired listeners. Ear and Hearing 2001;22:1–13. [PubMed: 11271971]

Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test. Vol. 3rd ed.. American Guidance Service; Circle Pines, MN: 1997.

Eisenberg LS, Martinez AS, Holowecky SR, Pogorelsky S. Recognition of lexically-controlled words and sentences by children with normal hearing and children with cochlear implants. Ear and Hearing 2002;23:450–462. [PubMed: 12411778]

Elliott LL, Clifton LAB, Servi DG. Word frequency effects for a closed-set word identification task. Audiology 1983;22:229–240. [PubMed: 6870665]

Field, A. Discovering statistics using SPSS. Vol. 2nd ed.. Sage; Thousand Oaks, CA: 2005.

Garlock VM, Walley AC, Metsala JL. Age of acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. Journal of Memory and Language 2001;45:468–492.

Goldman, R.; Fristoe, M. Goldman-Fristoe test of articulation. American Guidance Service; Circle Pines, MN: 2000.

Greenberg J, Jenkins C. Studies in the psychological correlates of the sound system of American English. Word 1964;20:157–177.

Hudgins C, Hawkins J, Karling J, Stevens S. The development of recorded auditory tests for measuring hearing loss of speech. Laryngoscope 1947;57:57–89.

Kirk, KI. Challenges in the clinical investigation of cochlear implant outcomes. In: Niparko, JK., et al., editors. Cochlear Implants: Principles and Practices. Lippincott Williams & Williams; Philadelphia: 2000. p. 225-259.

Kirk KI, Pisoni DB, Miyamoto RC. Effects of stimulus variability on speech perception in listeners with hearing impairment. Journal of Speech, Language, and Hearing Research 1997;40:1395–1405.

Kirk KI, Pisoni DB, Osberger MJ. Lexical effects on spoken word recognition by pediatric cochlear implant users. Ear and Hearing 1995;16:470–481. [PubMed: 8654902]

Kirk KI, Sehgal ST, Hay-McCutcheon M. Comparison of children's familiarity with tokens on the PBK, LNT, & MLNT. Annals of Otology, Rhinology, and Laryngology 2000;185:63–64.

Kucera, F.; Francis, W. Computational analysis of present-day American English. Brown University Press; Providence, RI: 1967.

Landauer TK, Streeter LA. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. Journal of Verbal Learning and Verbal Behavior 1973;12:119–131.

Levitt H. Transformed up-down methods in psychoacoustics. Journal of the Acoustical Society of America 1971;49:467–477. [PubMed: 5541744]

Logan, JS. A computational analysis of young children's lexicons. Indiana University; Bloomington, IN: 1992. (Research on Spoken Language Processing Technical Report No. 8)

Luce, PA. Neighborhoods of words in the mental lexicon. Indiana University; Bloomington, IN: 1986. (Research on Speech Perception Technical Report No. 6)

Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. Ear and Hearing 1998;19:1–36. [PubMed: 9504270]

Luce, PA.; Pisoni, DB.; Goldinger, SD. Similarity neighborhoods of spoken words. In: Altmann, G., editor. Cognitive models of speech perception: Pyscholinguistic and computational perspectives. MIT Press; Cambridge, MA: 1990. p. 122-147.

Mackersie CL. Tests of speech perception abilities. Current Opinions in Otolaryngology & Head and Neck Surgery 2002;10:392–397.

MacWhinney B, Snow C. The child language data exchange system. Journal of Child Language 1985;12:271–296. [PubMed: 4019604]

Mendel, LL.; Danhauer, JD. Historical review of speech perception assessment. In: Mendel, LL.; Danhauer, JL., editors. Audiologic evaluation and management and speech perception assessment. Singular Publishing Group; San Diego, CA: 1997. p. 1-5.

Miller GA, Heise GA, Lichten W. The intelligibility of speech as a function of the context of the test materials. Journal of Experimental Psychology 1951;41:329–335. [PubMed: 14861384]

Nusbaum, HC.; Pisoni, DB.; Davis, CK. Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. Indiana University; Bloomington, IN: 1984. (Research on Speech Perception Progress Report No. 10)

Pisoni DB, Nusbaum HC, Luce PA, Slowiaczek LM. Speech perception, word recognition and the structure of the lexicon. Speech Communication 1985;4:75–95.

Seewald RC, Ross M, Giolas TG, Yonovitz A. Primary modality for speech perception in children with normal and impaired hearing. Journal of Speech and Hearing Research 1985;28:36–46. [PubMed: 3981996]

Sommers MS. The structural organization of the mental lexicon and its contribution to age-related changes in spoken word recognition. Psychology and Aging 1996;11:333–341. [PubMed: 8795062]

Tice, R.; Carrell, T. Level16. Vol. Version 2.0.3. University of Nebraska; Lincoln, NE: 1998. [Computer software]

Tillberg I, Ronneberg J, Svard I, Ahlner B. Audio-visual speechreading in a group of hearing aid users: The effects of onset age, handicap age and degree of hearing loss. Scandinavian Audiology 1996;25:267–272. [PubMed: 8976000]

Treisman M. A theory of identification of complex stimuli with an application to word recognition. Psychological Review 1978a;85:525–570.

Treisman M. Space or lexicon? The word frequency effect and the error response frequency effect. Journal of Verbal Learning and Verbal Behavior 1978b;17:35–59.

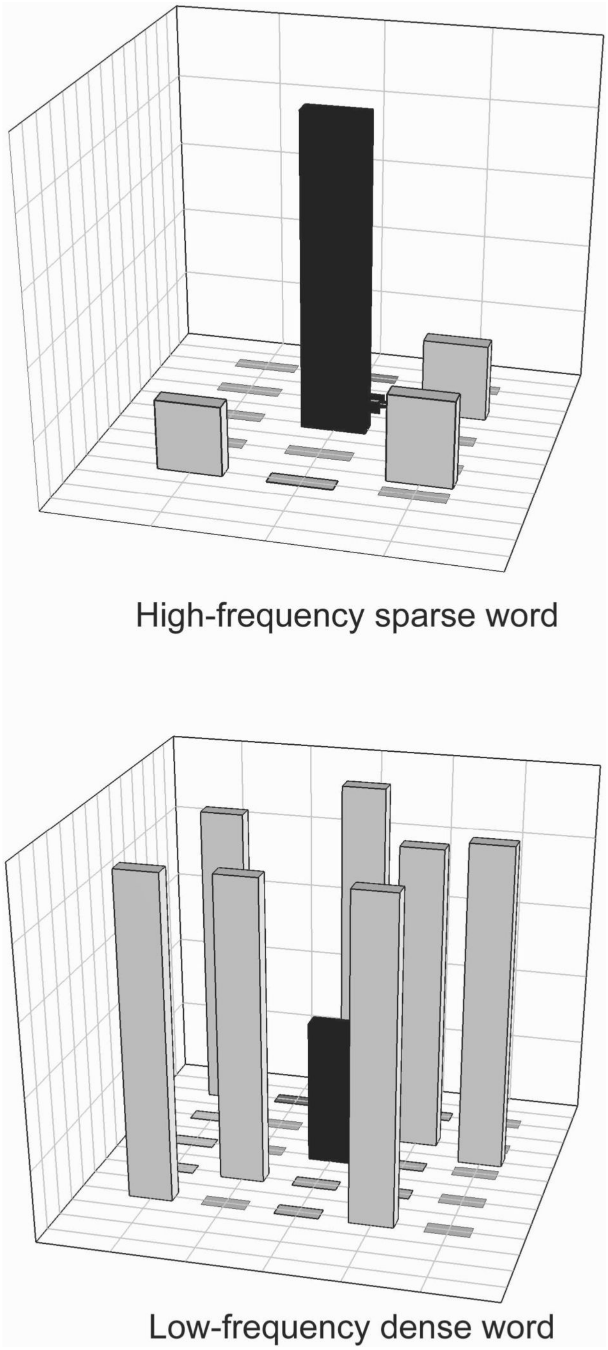

Figure 1.

Illustration of two hypothetical lexical neighborhoods. The target word is in black and its neighbors are in grey. Top: A high-frequency word belonging to a sparse neighborhood. Bottom: A low-frequency word belonging to a dense neighborhood.

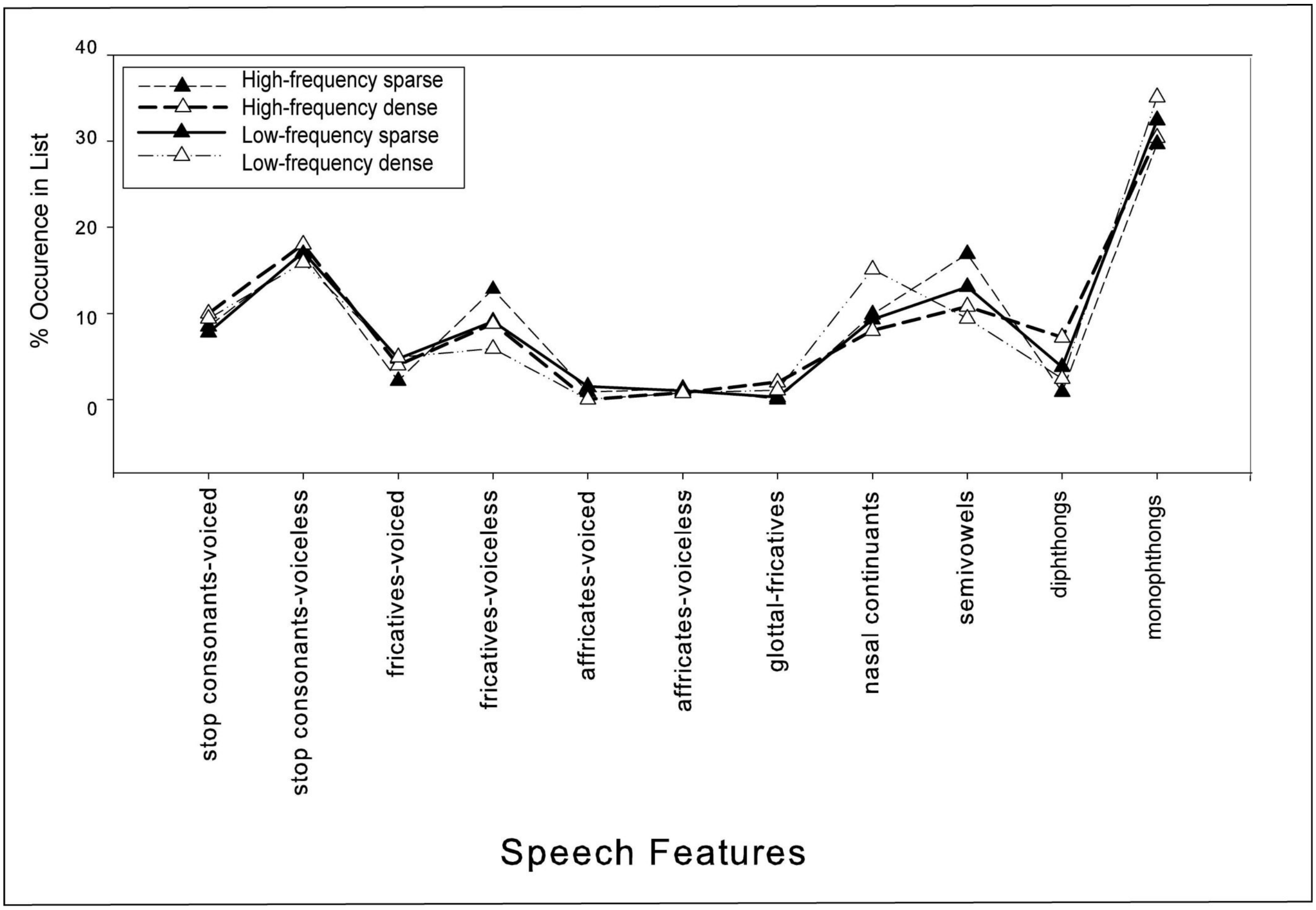

Figure 2.
The distribution of speech features by lexical category.

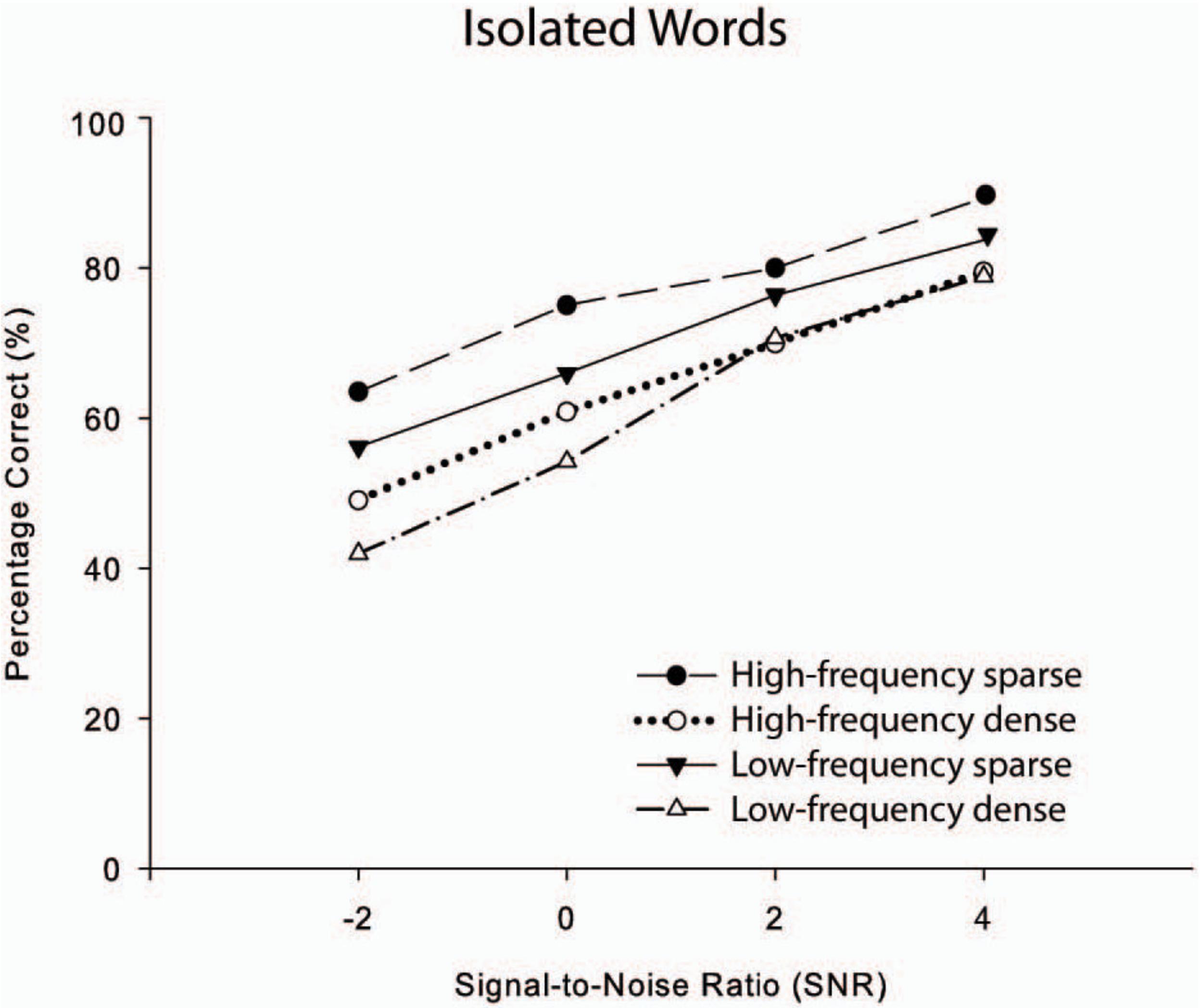

Figure 3.
Effect of SNR on isolated word recognition as a function of lexical category.

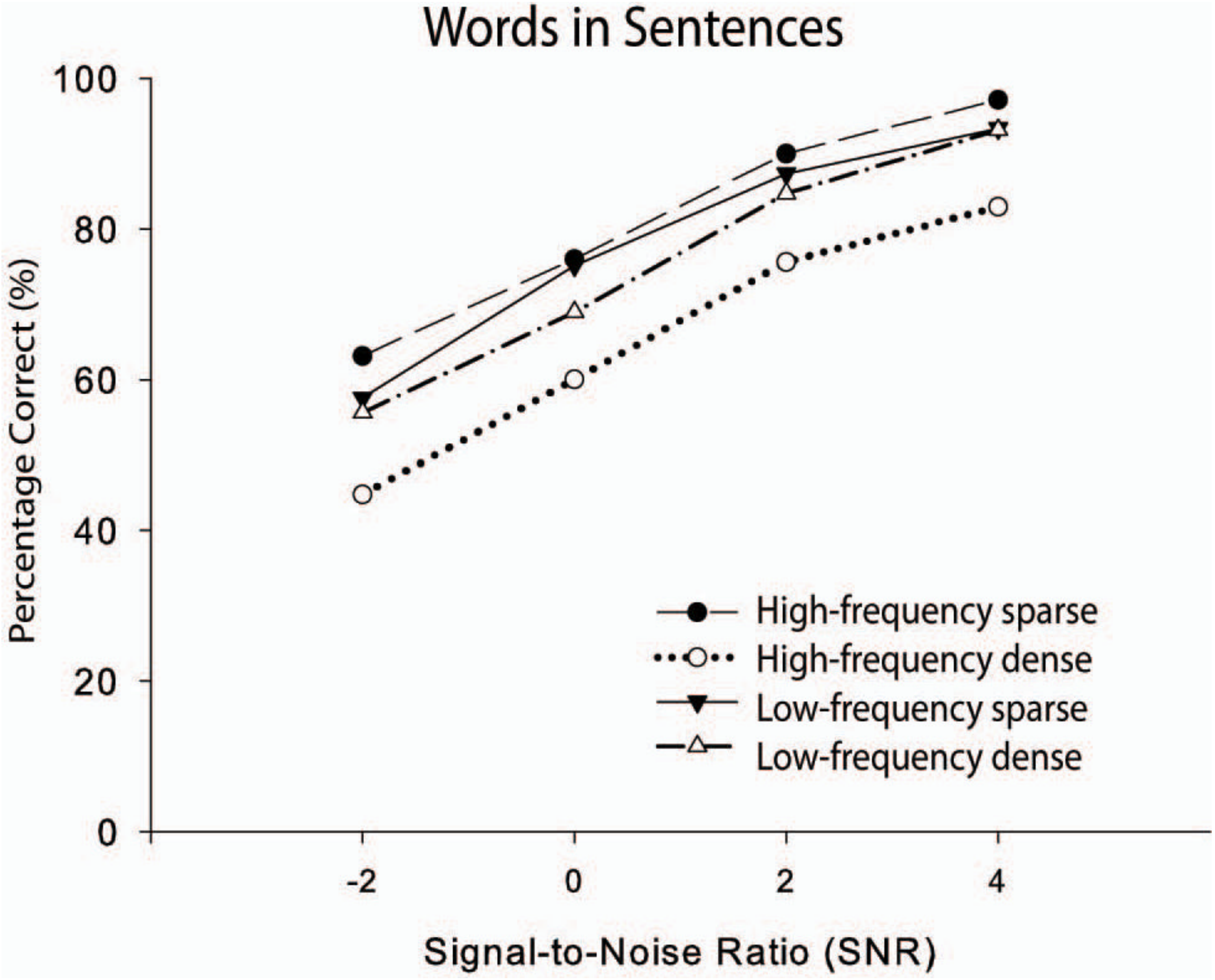

Figure 4.
Effect of SNR on recognition of key words in sentences as a function of lexical category.

Table 1

Lexical statistics for the corpus from which test items were drawn (Logan, 1992).

Description of Material	Word Frequency Median (Range)	Lexical Density Median (Range)
Monosyllabic Words [*]	4 (1–519)	4 (0–19)
Multisyllabic Words [+]	2 (1–100)	2 (0–7)
Monosyllabic and bisyllabic words [○]	1.75 (1–519)	1.25 (0–19)

[*] LNT created from this subset

[+] MLNT created from this subset

[○] Lexically-controlled sentences created from this subset

Table 2

Mean ratings of word familiarity across test measures. A seven-point scale was used where `1' represented a very unfamiliar word and `7' represented a very familiar word.

Dependant Variable	**Spoken Word Recognition Tests**				
	Lexical Neighborhood Test		**Multisyllabic Lexical Neighborhood Test**		**Phonetically Balanced Kindergarten Word Lists**
	High-frequency sparse words	Low-frequency dense words	High-frequency sparse words	Low-frequency dense words	
Familiarity Ratings (mean ± SD)	5.8 ± 1.6	5.6 ± 1.4	5.3 ± 1.4	5.0 ± 1.5	4.3 ± 1.6

Adapted from Kirk et al. (2000).

Table 3

Group demographics for participants in the main study. Standard deviation for the standard PPVT-III scores is shown in parentheses.

Demographic	**Signal-to-Noise Ratio (dB)**			
	−2	**0**	**2**	**4**
Mean Age (years: months)	9:2	8:10	9:1	9:1
Gender	3 boys,3 girls	3 boys,3 girls	4 boys,2 girls	3 boys,3 girls
Mean Standard PPVT-III Scores	118 (11)	115 (16)	114 (11)	112 (10)

Table 4

Mean lexical characteristics for the four classifications of sentences in the present investigation. Key words for the sentences were drawn from the Logan (1992) corpus.

Lexical Category	**Word Frequency** **Mean (Range)**	**Lexical Density** **Mean (Range)**
High-frequency sparse	8.45 (3–57)	0.63 (0–2)
High-frequency dense	10.51 (3–85)	5.48 (2–19)
Low-frequency sparse	1.25 (1–2)	0.27 (0–1)
Low-frequency dense	1.35 (1–2)	4.01 (2–13)

Table 5

SNRs (in dB) corresponding to a score of 30% and 70% correct in the sentence and word contexts.

Age-group	Sentences		Words	
	70%	30%	70%	30%
Group I: 5–7 year olds (n=5)	8	1	1	−4
Group II: 8–12 year olds (n=5)	2	−1	2	−3

Table 6

Isolated key word recognition as a function of lexical category and signal-to-noise ratio.

Lexical Category	Signal-to-Noise Ratio (dB)							
	−2		0		2		4	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
High-frequency sparse	64%	7	75%	5	80%	5	89%	5
High-frequency dense	49%	6	61 %	3	70%	6	80%	6
Low-frequency sparse	56%	6	66%	9	76%	4	84%	4
Low-frequency dense	42%	7	54%	7	71 %	8	79%	4

Table 7

Key word recognition in sentences as a function of lexical category and signal-to-noise ratio.

Lexical Category	**Signal-to-Noise Ratio (dB)**							
	−2		**0**		**2**		**4**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
High-frequency sparse	63%	8	76%	9	90%	3	97%	2
High-frequency dense	45%	9	60%	15	76%	7	83%	5
Low-frequency sparse	58%	9	75%	8	87%	4	93%	2
Low-frequency dense	56%	10	69%	10	85%	2	93%	4

Table 8

Correlations of mean word and sentence scores with lexical values of word frequency and lexical density for different signal-to-noise ratios.

Test Material	Lexical Characteristics	Signal-to-Noise Ratio (dB)			
		−2	**0**	**2**	**4**
Isolated Words (n=300)	Word Frequency	0.067	0.067	−0.019	−0.043
	Lexical Density	**−0.222**[*]	**−0.225**[*]	**−0.239**[*]	**−0.188**[*]
Sentences (n=100)	Word Frequency	0.009	0.006	−0.004	−0.05
	Lexical Density	**−0.297**[*]	**−0.335**[*]	**−0.339**[*]	**−0.423**[*]

[*] $p<0.01$