An examination of word frequency and neighborhood density in the development of spoken-word recognition

JAMIE L. METSALA
University of Maryland, College Park, Maryland

In this study, the effects of word-frequency and phonological similarity relations in the development of spoken-word recognition were examined. Seven-, 9-, and 11-year-olds and adults listened to increasingly longer segments of high- and low-frequency monosyllabic words with many or few word neighbors that sounded similar (neighborhood density). Older children and adults required less of the acoustic–phonetic information to recognize words with few neighbors and low-frequency words than did younger children. Adults recognized high-frequency words with few neighbors on the basis of less input than did all three of the children's groups. All subjects showed a higher proportion of different-word guesses for words with many versus few neighbors. A frequency × neighborhood density interaction revealed that recognition is facilitated for high-frequency words with few versus many neighbors; the opposite was found for low-frequency words. Results are placed within a developmental framework on the emergence of the phoneme as a unit in perceptual processing.

Investigations of adult speech recognition have focused on the form of internal lexical representations and organizational properties of the mental lexicon. Many investigations have also been directed toward delineating the perceptual abilities of infants. There has been much less attention given to developmental issues in spoken-word recognition from early through middle childhood (Walley, 1993a). There are, however, compelling theoretical reasons to propose that children's spoken-word representations may not be the same as those of adults. There are also applied reasons for investigating developmental changes from early to middle childhood. Due to the size and stability of children's vocabulary, word-recognition processes may differ from those of adults, and these developmental changes may have important effects on the emergence of phonemic awareness and reading.

From Wholes to Segments: Lexical Restructuring

Phonemic representations form the basis of most adult models of spoken-word recognition, and therefore represent the goal state of lexical organization (for a review, see Pisoni & Luce, 1987). Phonemes may form the basis of lexical representations initially or, alternatively, the phoneme may emerge as a perceptual processing unit only

with increased vocabulary growth. The view that initial word recognition is relatively holistic and that representations become more segmentally based with development is gaining prevalence (Jusczyk, 1986, 1992; Walley, 1993b).[1] Aslin and Smith (1988) have suggested that since words form the fundamental unit of meaning, the word may be the initial level of linguistic contrast (see, also, Walley, Michela, & Flege, 1994). That is, children beginning to talk are preoccupied with isolating words from the speech stream and building correspondences between words and their meanings. It may be that only with vocabulary growth and the resulting need to discriminate between a growing number of alternative words in long-term memory that lexical representations are forced to become segmentally (phonemically) encoded.

My suggestion, therefore, is that the developmental process of speech perception involves learning to perceive stimulus words as complexes of segments (see also Walley, 1993b). This process results in the establishment of new, restructured representations that support spoken-word recognition.[2] The change from relatively holistic to more segmentally based lexical representations has been proposed to be a protracted process that may extend into early childhood (Fowler, 1991; Walley, 1993b). Such developmental changes may have important consequences for some perceptual and cognitive abilities that develop beyond early childhood. For example, Fowler (1991) has suggested that children's emerging phonemic awareness skills from around 5 to 8 years of age may be related to developmental changes in lexical representations. Since phonemic awareness ability is a potent predictor of reading achievement (e.g., Bradley & Bryant, 1978, 1991; Cunningham, 1990; Lundberg, 1991; Lundberg, Frost, & Peterson, 1988; Treiman & Breaux, 1982; Treiman &

Zukowski, 1991; Wagner, Torgeson, Laughon, Simmons, & Rashotte, 1993; Wagner, Torgeson, & Rashotte, 1994), one priority is to understand factors influencing the emergence of the phoneme at a perceptual level.

A second processing advantage afforded by segmental representations will be at the perceptual level, namely, spoken-word recognition based on partial acoustic–phonetic information. According to adult models of speech perception (e.g., Luce, 1986; Pisoni & Luce, 1987), acoustic–phonetic input is mapped onto segmental representations in an on-line manner, and words therefore can be recognized before their acoustic offset. Alternatively, more holistic representations will map onto larger incoming linguistic units (e.g., syllables or words), and therefore recognition will be relatively delayed. A goal of the present study was to empirically examine factors hypothesized to be important in the developmental time course of lexical–phonemic restructuring and the trend toward less holistic word-recognition processes.

Empirical Evidence Bearing on Children's Spoken-Word Recognition

There has been little evidence bearing on spoken-word recognition in early to middle childhood, and few developmental investigations have systematically examined the effects of word frequency on recognition (e.g., Cole, 1973, 1981; Cole & Perfetti, 1980; Elliott, Hammer, & Evan, 1987; Walley, 1987, 1988; but see Elliott, Clifton, & Servi, 1983, and Fox & Koenigsknecht, 1989). So far, it appears that children do engage in more holistic processing of spoken words than do adults. Walley (1988) found that adults were most disrupted by noise-replaced segments in the beginnings of words. Five-year-olds were less sensitive to the position of a replaced segment, their attention being more evenly distributed across the entire word. Similarly, Walley, Smith, and Jusczyk (1986, Experiment 1) found that 8-year-olds, but not 5-year-olds, were able to classify nonwords together on the basis of a single common phoneme; younger children needed similarity across more of the acoustic–phonetic patterns in order to classify nonwords together (see, also, Treiman & Breaux, 1982).

More holistic lexical representations are also suggested by the findings that 3-, 4-, and 5-year-old children's productions and perceptions of CVC syllables and individual phonemes are more influenced by coarticulatory factors than is perception in 7-year-olds and adults (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, Studdert-Kennedy, & McGowan, 1989). Also, children have been found to need more of the bottom-up input from word onset than do adults in order to recognize even very familiar words (Elliott et al., 1987; Fox & Koenigsknecht, 1989; Walley, 1988). The greater amount of input required, despite children's smaller lexicons, suggests that input maps onto units in children's lexicons that are larger than those in adults, and that children are not as able to use word-partial acoustic–phonetic input to support word recognition.

The Role of Vocabulary Growth in Lexical Restructuring

An issue of primary importance is what drives this shift from relatively holistic to more segmentally based processing of spoken words. One possibility is that increased vocabulary size, accompanied by the need to discriminate among a growing number of similar-sounding alternatives, causes segmental restructuring. For example, recognizing the word "big" will require a more detailed representation if the child's vocabulary also includes such items as bag, bug, bit, dig, and wig than if it does not. Segmental restructuring is not hypothesized to occur in an all-or-none, system-wide fashion, but rather on an item-by-item basis. At any point in development, the degree of segmentation for a given lexical item may depend on the number of similar-sounding neighbors; thus, the extent of segmentalized representations will vary between lexical items.

Alternative, more holistic lexical representations for young children are congruent with analyses of computerized lexicons. Computational analyses suggest that words in 5-year-olds' lexicons are more structurally distinct than those of 7-year-olds, which in turn are more distinct than the lexicons of adults (Charles-Luce & Luce, 1990, 1995; but see Dollaghan, 1994). Logan (1992) examined the computational viability of systems alternative to those of phonemic representations in on-line lexicons modeled after those of children 1½ to 5 years of age. Neighborhood densities calculated on the basis of manner class representations were very similar to those based on phonemes and, he suggested, could be one feasible alternative that might underlie young children's spoken-word recognition.

The primary purpose of the current study was to examine, empirically, the role of structural similarity relations in the developmental trend toward less holistic processes in spoken-word recognition. Age-related changes in spoken-word recognition were examined across middle childhood (ages: 7, 9, and 11 years) and compared with adult performance as a function of phonemic similarity relations and word frequency. The best performance on the present perceptual task (gating) should occur for words that the subject does not perceive holistically but as a complex of smaller linguistic segments. I therefore predicted a developmental decrease in the amount of bottom-up input needed to recognize words, especially those with few similar-sounding neighbors. Although it is not surprising to predict that younger children will do more poorly than older children and adults, given their more structurally distinct lexicons, some models would predict that children should need less information than adults.[3]

While it has been suggested that phonemic similarity relations play a pivotal role in the developmental time course of segmental restructuring for individual lexical items (Jusczyk, 1992; Walley, 1993b), there is no empirical evidence for such. The theory is that words that have many similar-sounding neighbors will undergo, develop-

mentally early, segmental restructuring. That is, these words need to be encoded phonemically at an early age in order to be discriminated from similar-sounding words in the listener's lexicon. Words with many neighbors should therefore be recognized on the basis of less bottom-up input than words that are stored relatively more holistically or have undergone segmental restructuring more recently. Similarly, due to increased experience with the sound pattern, frequency may also play a causal role in representations becoming more segmentalized. Walley and Metsala (1990) found developmental differences among 5- and 8-year-olds and adults on mispronunciation judgments for words that were rated to be acquired relatively late in development, but no such differences were found for the earliest acquired words. In the present study, the prediction was that developmental differences would be most pronounced for words that had few similar-sounding neighbors and were not heard frequently—those words that would be chronologically latest to undergo segmental restructuring.

The combined effects of word frequency and neighborhood density are not easy to predict. First, the required empirical investigations have not been carried out with children. Second, there are inconsistencies in the relevant adult investigations. Luce (1986) found that adults were quicker at an auditory naming task for words with few as opposed to many similar-sounding neighbors. On a lexical decision task, however, subjects showed an advantage for low-frequency words with many as opposed to few neighbors. A similar interaction for frequency and neighborhood density has been found in studies of adults' word reading (e.g., Andrews, 1992; Brown, 1987; Luce, 1986; Seidenberg & McClelland, 1989). A second focus of this study, therefore, was to elucidate the combined effects of neighborhood density and frequency in spoken-word recognition for children and adults.

The Gating Task

The gating paradigm was used to investigate listener and stimulus characteristics on spoken-word recognition. In this paradigm, listeners are presented with increasing amounts of acoustic–phonetic information, from word onset over a series of trials, and asked to identify the target after each gate (Grosjean, 1980). This is an off-line task insofar as listeners respond to the stimulus after its offset and there is no response deadline. Investigations that have used the gating task with adults have replicated such important aspects of speech processing as frequency, word length, and context effects (Grosjean, 1985; Salasoo & Pisoni, 1985); furthermore, the amount of information needed to isolate a word in the gating task, in absolute terms, is similar to recognition times in on-line tasks such as speech shadowing and word monitoring (e.g., Marslen-Wilson, 1984, 1987). Studies have also shown that the absence of a response deadline does not enhance performance (Tyler & Wessels, 1985), nor does the successive presentation format (Cotton & Grosjean, 1984; Salasoo & Pisoni, 1985).

The gating task has been increasingly used to examine the speech recognition of both children (e.g., Elliott et al., 1987; Fox & Koenigsknecht, 1989; Walley, Michela, & Wood, 1995) and atypical populations (e.g., Elliott, Scholl, Grant, & Hammer, 1990; Metsala, in press). In one investigation, Walley et al. (1995) found that children and adults responded to different presentation formats in a qualitatively similar manner. These investigators concluded that the gating paradigm was suited to developmental comparisons. Indeed, the off-line characteristics of this task make it especially suitable for young listeners as there is no imposed reaction time deadline, yet there is precise control over the amount of bottom-up input that the listener's response is based upon.

In summary, the primary purpose of the present study was to examine developmental changes in spoken-word recognition. The effects of neighborhood density on children's recognition performance have not been previously examined, nor have the combined effects of word frequency and neighborhood density. The present study allowed for an examination of these two variables in children's and adults' word recognition, as well as developmental comparisons.

METHOD

Subjects

The subjects were 20 first- and second-grade children ($M =$ 7 years 3 months, range = 6 years 4 months to 8 years 4 months; 13 girls, 7 boys), 20 third- and fourth-grade children ($M = 9$ years 2 months, range = 8 years 5 months to 10 years; 11 girls, 9 boys), 20 fifth-grade children ($M = 10$ years 10 months, range = 10 years 1 month to 11 years 6 months; 8 girls and 12 boys), and 18 graduate students (range = 20–40 years; 10 females and 8 males). The children were recruited from two predominantly middle-class public schools and were reading above the 25th percentile on the Wide Range Achievement Test–Rev. All subjects were native speakers of English and had no reported history of speech or hearing problems. Three additional subjects from the youngest group did not complete the gating task because of absence or inattention.

Stimuli

The 28 target words, all concrete nouns, formed a 2 (word frequency) × 2 (neighborhood density) stimulus design (see Table 1). Half of the words were of high frequency and half were of low frequency, according to Kučera and Francis (1967). Neighborhood density was defined as the number of words that differed from the target word by a one-phoneme addition, deletion, or substitution (see Charles-Luce & Luce, 1990; Goldinger, Luce, & Pisoni, 1989; Luce, 1986). Half of the high- and low-frequency words had many neighbors (dense neighborhoods), and half had few neighbors (sparse neighborhoods). These cells were matched on the mean log frequency of a target-word's neighbors.

For target selection and neighborhood statistics, I started with Luce's (1986) database of 918 words previously used in perceptual experiments with adults. These were monosyllabic words from Webster's dictionary which had received, from adults, a 6 or above on a 7-point familiarity rating scale (Nusbaum, Pisoni, & Davis, 1984). To ensure that target words and neighborhood statistics would be developmentally appropriate, age-of-acquisition (AOA) ratings were obtained. Sixty-four adult volunteers each rated about one quarter of the 918 words on a 9-point AOA scale in the manner described by Carroll and White (1973). On this scale, a rating

Table 1
Word Stimulus Characteristics: Mean Log Frequency, Number of Neighbors, and Mean Log Frequency of Neighbors

Word Type	Mean Log Frequency of Target	Mean Number of Neighbors	Mean Log Frequency of Neighbors
High–dense	3.16	14.86	2.22
High–sparse	3.05	4.43	2.13
Low–dense	1.28	12.57	2.34
Low–sparse	1.34	3.29	2.15

Note—Word types *high* and *low* refer to words of high- versus low-frequency stimulus cells; *dense* and *sparse* refer to neighborhood structure: words from dense neighborhoods have many similar-sounding words in the listener's lexicon as opposed to sparse words, which have few similar-sounding words in the listener's lexicon. Frequency counts are based on Kučera and Francis (1967).

of 1 = *acquired at age 2*; 5 = *acquired at age 6*; and 9 = *acquired at age 13 or above*. Words that received a mean rating below 7 (*acquired before about 9 years of age*) formed the lexicon from which neighborhood statistics were calculated. All of the 28 target words had mean AOA ratings that corresponded to being acquired at or before 7 years of age, and the mean AOA rating for all four stimulus cells corresponded to being acquired at or before 5–6 years of age. Previous research has shown that AOA ratings are reliable and are valid predictors of the vocabulary knowledge of children (Brown & Watson, 1987; Carroll & White, 1973; Gilhooly, 1984; Gilhooly & Gilhooly, 1980; Gilhooly & Logie, 1982; Walley & Metsala, 1990, 1992).

The target words, as well as four practice items, were read by a male speaker and recorded in a sound-attenuated booth, then band-pass-filtered (60 Hz–10 kHz) and digitized at a 10-kHz sampling rate with 12-bit amplitude resolution. The test words were excised, normalized for peak amplitude, and stored on the computer. Gated versions of the words were prepared using a program for truncating digitized stimuli. The first gate for each word was 100 msec in duration from word onset, and each subsequent gate was an additional 50 msec from word onset (e.g., Gate 2 = 150-msec duration; Gate 3 = 200-msec duration, etc.), until the last trial, on which the complete word was presented. The test stimuli were converted from digital to analog form and recorded in two random orders to form Tapes A and B. Each tape had the same four gated words for the practice trials.

Table 1 presents the log frequency, number of neighbors, and mean neighbor log frequency for each of the frequency × neighborhood density cells (see the Appendix for similar information for each of the 28 words). The number of gates per word as well as the mean log frequency of the neighbors were submitted to two separate analyses of variance (ANOVAs), with frequency and density as between-item factors. The analyses indicated that there were no differences between the four cells for mean neighbor log frequency. Importantly, the four cells also did not differ in terms of mean number of gates per stimulus item.

Procedure

The subjects were tested individually in a quiet room in their school, where they listened to one of the two random orders of the taped stimuli. All subjects were told that they would first hear only the very beginning of a word and then, on each trial, would hear "more and more" of that word; after each presentation, the subjects guessed what they thought the word was. The subjects first completed the four practice words, and even the youngest children were able to complete the task after the first practice item.

The gated stimuli were presented over headphones at a comfortable listening level via a Sony cassette tape player. After each stimulus, the subject's response was transcribed by the experimenter. If the subject did not respond during the interstimulus interval, then a blank response was recorded.[4] Even the youngest

subjects became sensitive to this interstimulus interval during the practice trials, and offered their responses well within this time limit. No contingent feedback was given during the gating test, but general encouragement was given to all the subjects during testing. The subjects completed the gating task in one session lasting approximately 30 min.

RESULTS

Isolation Points

The isolation point (IP) was the stimulus duration at which a subject first correctly guessed a target word (consistent with Walley et al.'s definition of IP).[5] For target words that were not correctly identified, the IP was calculated to be the total duration of the word plus 50 msec, the duration of one gate (see also Elliott et al., 1987; Walley et al., 1995).[6] All effects reported below were significant at $p < .05$, unless otherwise stated.

Subjects' mean IPs were submitted to a 4 (age) × 2 (frequency) × 2 (neighborhood density) ANOVA for a mixed design. Table 2 shows the mean IP duration and standard deviations for each cell of the analysis. The analysis revealed main effects for age [$F(3,74) = 18.24, p < .0001$] and for frequency and density [$F(1,74) = 951.16$ and $10.85, p < .005$], as well as a frequency × neighborhood density interaction [$F(1,74) = 241.08, p < .0001$] and an age × frequency × neighborhood density interaction [$F(3,74) = 2.92$].[7] The three-way interaction revealed different patterns of developmental differences across the four cells (see Figure 1). The two youngest groups of children identified low-frequency (LF) words at longer gate durations than did older children and adults. For high-frequency (HF) words in sparse neighborhoods (HF–sparse), 7-year-olds needed more input for recognition than did 11-year-olds, and all three children's groups needed more input than did adults ($M = 200$ vs. 234, 256, and 274 msec, $SD = 23.11, 36.72, 39.59$, and 38.32, for adults to the youngest group, respectively). For HF words in dense neighborhoods (HF–dense), there were no significant age differences. Overall HF words had shorter gate durations than LF words ($M = 275$ msec vs. 385 msec, $SD = 34.25$ and 34.88).

All age groups displayed a frequency × neighborhood density interaction; recognition of HF words was facilitated for sparse versus dense neighborhoods ($M = 242$ msec

Table 2
Mean Isolation Point (in Milliseconds) as a Function of Age,
Frequency, and Neighborhood Density

Group Ages	Word Type							
	High–Dense		High–Sparse		Low–Dense		Low–Sparse	
	M	SD	M	SD	M	SD	M	SD
7 years	321.79	39.63	274.29	38.31	390.00	32.56	426.07	40.95
9 years	318.93	39.00	256.43	39.59	380.36	38.83	426.43	34.22
11 years	302.14	38.92	233.57	36.72	345.36	42.62	386.43	38.56
Adults	290.08	22.87	200.00	23.11	337.70	39.18	390.87	17.46

Note—Word types *high* and *low* refer to words of high- versus low-frequency stimulus cells; *dense* and *sparse* refer to neighborhood structure: words from dense neighborhoods have many similar-sounding words in the listener's lexicon as opposed to sparse words, which have few similar-sounding words in the listener's lexicon. Frequency counts are based on Kučera and Francis (1967).

vs. 309 msec, $SD = 44.22$ and 37.63). The opposite was observed for LF words for which recognition was based on less input in dense versus sparse neighborhoods ($M = 364$ msec vs. 408 msec, $SD = 43.80$ and 38.73; see Figure 1).

Item Analyses

For each age group, an ANOVA with word items as the random effect and frequency and density as the between-item factors was completed. For the adult group, the ANOVA showed a main effect of frequency [$F(1,24) = 17.79, p < .001$] and a frequency × neighborhood density interaction [$F(1,24) = 6.42, p < .02$]. This parallels the subject analysis for adults. For the oldest group of children, there was a main effect of frequency [$F(1,24) = 9.48, p < .01$]; the frequency × neighborhood density in-

teraction did not reach traditional levels of significance [$F(1,24) = 2.97, p < .10$]. This item analysis parallels the subject analysis for this age group with the exception that the frequency × neighborhood density interaction was significant for the subject analysis. Both item analyses for the two youngest groups of children showed main effects of frequency only [$Fs(1,24) = 11.87$ and 11.16, $p < .01$, for the middle and youngest groups, respectively]. For these two youngest groups of subjects, the frequency × neighborhood density interaction, which was significant for the subject analysis, was not significant for the item analysis.

Initial Phoneme

The trial on which the subjects first guessed a word candidate with the correct initial phoneme was calcu-
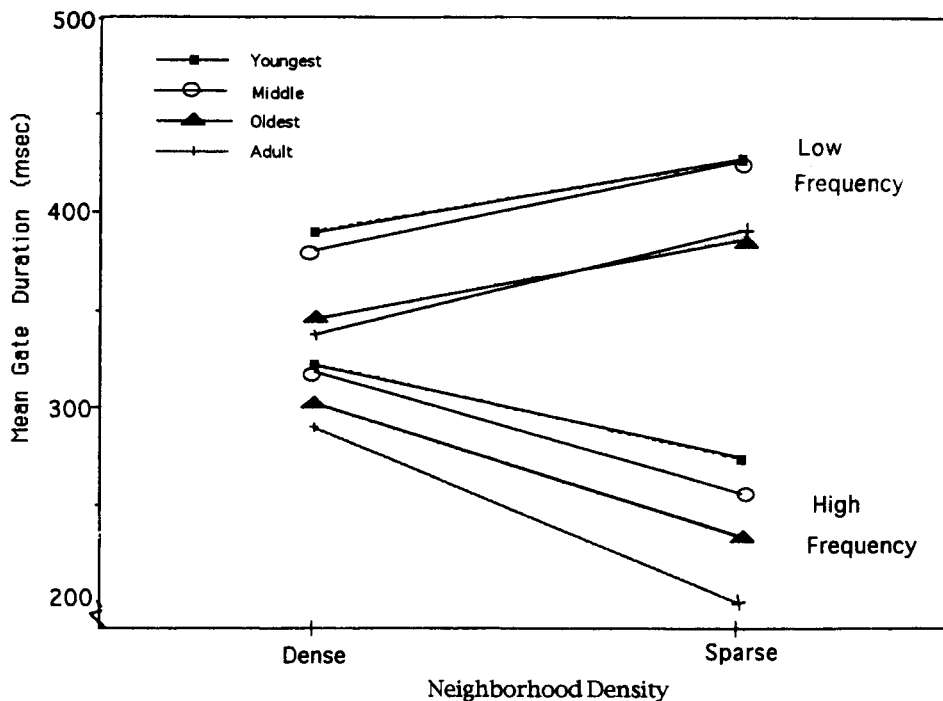

Figure 1. Mean isolation point (in milliseconds) as a function of age, frequency, and neighborhood density.

lated, and an ANOVA for the factor of age was significant [$F(3,74) = 12.11, p < .0005$]. The youngest group of children guessed a word with the correct initial phoneme on a later trial than did the oldest group of children and adults (M gate number = 2.46 vs. 1.64 and 1.37, SD = .68, .47, and .40; corresponding to mean durations of 173 msec vs. 131 and 118 msec). The middle group of children also guessed a word candidate with the correct initial phoneme on a later trial than did adults (M trial = 1.99 vs. 1.37, SD = .75 and .40).

Word Candidates

The total recognition trial (TRT) was defined as the gate for which a subject correctly identified the target word without subsequently changing her or his response. For each subject, the number of different incorrect guesses was then calculated and divided by the number of pre-TRT trials. This is a measure of the proportion of different word candidates generated prior to the subject's maintaining the correct response (Walley et al., 1995).

The 4 (age) × 2 (frequency) × 2 (neighborhood density) ANOVA revealed a main effect of age [$F(3,74)$ = 4.76, $p < .005$]. Adults had a higher proportion of different guesses than did the two youngest groups of subjects (M proportion = .56 vs. .45 and .48, SD = .10, .11, and .12). There was also a main effect of density [$F(1,74)$ = 11.34, $p < .005$]. There was a higher proportion of different guesses for dense neighborhoods than for sparse neighborhoods (M proportion = .52 vs. .49, SD = .12 and .12). Table 3 shows the mean proportion of different guesses prior to recognition for each cell of the ANOVA.

DISCUSSION

My findings support the hypothesis that the development away from more holistic processing of spoken words is related to the sound-similarity relations and experienced frequency of individual lexical items. Developmental comparisons showed that with increasing age, less acoustic–phonetic information was needed to recognize high-frequency words in sparse neighborhoods, as well as low-frequency words in both sparse and dense neighborhoods. It appears that children learn to represent the

segmental distinctions that are necessary for discrimination among lexical alternatives. The extent and developmental time course of lexical restructuring appears to be a function of a word's location (in terms of neighborhood structure) in the child's mental lexicon.

The increasing shift from relatively holistic to more segmental processing is, I suggest, the result of the increasing segmental structure of lexical representations. The impetus for such restructuring is growing lexical knowledge—the increasing need to discriminate among many similar-sounding alternatives and the increased exposure of individual lexical items. Developmental differences were pronounced for words that were not heard frequently and for words that did not have many similar-sounding neighbors. That is, developmental differences were found for those words that, I proposed, are more holistically represented and are the latest, chronologically, to move toward segmental restructuring. Young children were able to recognize the words that, I suggest, undergo segmental restructuring earliest—that is, high-frequency words with many similar-sounding neighbors—with as little of the bottom-up information as adult listeners had.

Developmental Changes Linked to Lexical Representations

I believe that the observed pattern of results links development in spoken-word recognition to changes in lexical representations to a greater extent than previous research has shown. First, alternative processing explanations are not sufficiently powerful to explain my findings. For example, an alternative account of these developmental changes is that children *are* processing words in a segmental manner, but strategically pay equal attention to segments throughout a word—possibly because young listeners are more likely to encounter new, unknown words than adults are. If this alternative explanation was correct, however, there should have been consistent developmental differences across all of the stimulus cells. Even a stronger version of this argument, that this bias toward paying attention to information throughout a word is dominant only when children are less familiar with a word (e.g., low-frequency words), does not account for the present findings. Young children also

Table 3
Proportion of Different Word Guesses Prior to Total Recognition Trials

Group Ages	Word Type							
	High–Dense		High–Sparse		Low–Dense		Low–Sparse	
	M	SD	M	SD	M	SD	M	SD
7 years	.468	.138	.412	.134	.477	.141	.429	.117
9 years	.480	.142	.489	.203	.483	.115	.456	.113
11 years	.568	.133	.516	.126	.509	.097	.503	.094
Adults	.604	.128	.575	.154	.554	.110	.524	.100

Note—Word types *high* and *low* refer to words of high- versus low-frequency stimulus cells; *dense* and *sparse* refer to neighborhood structure: words from dense neighborhoods have many similar-sounding words in the listener's lexicon as opposed to sparse words, which have few similar-sounding words in the listener's lexicon. Frequency counts are based on Kučera and Francis (1967).

needed more information for high-frequency words in sparse neighborhoods—words with which they were very familiar, but for which there was no need to discriminate from many similar-sounding alternatives.

A second reason for suggesting that the current findings support the notion that developmental changes are linked to lexical representations relates to the observation that children required more information than adults did in order to guess a word with the correct initial phoneme as the target word. This is consistent with the hypothesis that representations become more segmental with lexical growth. Models of mature word recognition propose that incoming acoustic–phonetic information is matched, in an on-line manner, to phonemically based lexical representations (for a review, see Pisoni & Luce, 1987). The more holistic representations underlying children's spoken-word recognition would require more of the bottom-up input in order for words with the correct initial phoneme to become active candidates. My observations are consistent with studies that have found that children's ability to perceive and produce CVC syllables is more influenced by coarticulatory factors than is that of adults (Nittrouer & Studdert-Kennedy, 1987; Nittrouer et al., 1989).

On the basis of the results, I suggest that there are continued perceptual advantages linked to changes in lexical representations past early childhood. These developmental findings support the thesis that the phoneme emerges first as a perceptual processing unit with vocabulary growth, rather than being an inherent unit in the organization of the mental lexicon. The interaction of age × frequency × neighborhood density supports the view that segmental restructuring occurs in a word-by-word manner rather than on a system-wide basis. That is, developmental differences in recognition vary with the extent of segmentalized representations across different locations in the mental lexicon. In addition, it appears that both frequency and neighborhood density are important determinants of the developmental time course toward phonemic restructuring for individual items.

The Frequency × Neighborhood Density Interaction: A Developmental Perspective

Previous studies have not been consistent in their findings of the combined effects of frequency and neighborhood density on adults' spoken-word recognition. The present results are consistent with subjects' performance on an auditory lexical decision task (Luce, 1986) and with studies on visual word recognition (Andrews, 1992; Seidenberg & McClelland, 1989). Recognition of high-frequency words was facilitated for sparse versus dense neighborhoods; low-frequency words were recognized at shorter gate durations in the dense versus sparse neighborhood condition.

The observed frequency × neighborhood density interaction may result from inconsistencies in two functionally distinct influences—the results of which are consistent with the proposed developmental theory of the time course of lexical restructuring. These two distinct influences are most apparent in models of cognitive processes that instantiate distributed representations (e.g., see Seidenberg & McClelland's, 1989, model of visual word recognition). These will be called on-line effects and structural-residual effects. On-line effects are defined as those that result from competition among items that most closely match the bottom-up input. Structural-residual effects are defined as those that arise from the structural properties of lexical representations and result from the developmental (or acquisition) phase—therefore, residual. These are effects on recognition due to characteristics of lexical representations (e.g., segmentalized) that result from the history of the individual word in the mental lexicon.

With increasing neighborhood density, the effects from these two functionally distinct influences will be opposite. Structural-residual effects will facilitate recognition as neighborhood density increases in two ways. First, there will be pressure on holistic representations to undergo segmental restructuring in order to discriminate between many similar-sounding words. Second, sublexical patterns of activation will be more often instantiated for words with many similar-sounding neighbors, resulting in greater connection strengths between distributed units (e.g., see Seidenberg & McClelland, 1989). On the other hand, on-line effects in neighborhoods of increasing density will impede recognition. Competition for any given acoustic–phonetic pattern will depend on the number and frequency of word neighbors, as well as on the frequency of the target word itself (Luce, 1986).

Structural-residual and on-line effects will act consistently by enhancing recognition of words of increasing frequency. Words that are heard often will have a greater opportunity to result in complete, segmental representations (structural-residual effect) and will be afforded priority in the competitive process (e.g., through activation weights or decision bias—the on-line effect).

Thus, effects associated with increasing frequency are consistent across the two functionally distinct influences. For words of increasing neighborhood density, however, the structural-residual effects enhance recognition, whereas on-line effects impede recognition. I suggest that the observed frequency × neighborhood density interaction is a result of these combined effects. Specifically, for low-frequency words, the structural-residual effects of neighborhood density are primary. Low-frequency words in dense (but not sparse) neighborhoods will have the segmentalized structure to support early recognition and are therefore recognized with less input than similar words in sparse neighborhoods. Frequently heard words have the opportunity to become segmentally based over time, regardless of neighborhood structure. For these words, on-line competition is the determining factor in recognition performance, recognition therefore being facilitated for high-frequency words in sparse as opposed to dense neighborhoods. For children even younger than those tested in the present study, the prediction would be

that all words would be of relatively lower frequency and that neighborhood density might be the initial impetus for segmental restructuring of lexical items.

The proposal that recognition is facilitated by the experienced frequency of sublexical units in spoken-word recognition is not unprecedented. Landauer and Streeter (1973) found that high-frequency, monosyllabic words that contained more phonemes found in common words were recognized better, when presented in noise, than were words containing phonemes typical of rare words. Consistent with Landauer and Streeter's findings and with models of visual word recognition, my proposal is that increased similarity neighborhoods have a positive effect on the recognition of spoken words, in part due to the frequency of sublexical patterns of phonemes. Brown and Watson (1987) have proposed that processing advantages for early-acquired words are due to more robust representations. The current paper suggests how both frequency and neighborhood density may interact to afford differential processing advantages for words in the mental lexicon, thus operationalizing the notion of "more robust" representations.

The finding of a higher proportion of different word candidates generated prior to recognition in dense versus sparse neighborhoods also provides support for the importance of word neighborhoods in spoken-word recognition (see also Charles-Luce & Luce, 1995; Dollaghan, 1994; Goldinger et al., 1989; Logan, 1992; Luce, 1986). This finding was similar for children and adults, validating the use of an on-line lexicon, modeled after the vocabulary knowledge of words estimated to be known by an age just above that of my youngest subjects, for making developmental comparisons.

One limitation of the present study is that neighborhood statistics calculated on the basis of a computerized lexicon are only an estimate for individual subjects (Luce, 1986). The frequency × neighborhood density interaction may have been apparent only in the item analysis for older children and adults due to this limitation. Younger children are still learning many new words, and the density status of any one word may vary more between subjects in the youngest age groups. Future research with young children might examine recognition performance of target words for which neighborhood structure is empirically determined on a subject-by-subject basis.

Summary

Studies with children 5 years of age and up have focused more on phonemic awareness abilities than on basic word-recognition processes. An implicit assumption of these studies has been that lexical items are fully segmented much earlier than the age at which access of these units is necessary for reading (for discussion, see Fowler, 1991). However, in order to understand the development of, and individual differences in, phonemic awareness skills and subsequent reading ability, the emergence of the phoneme as a basic unit in spoken-word recognition needs to be delineated (Fowler, 1991; Walley, 1993b). The present study takes an empirical step toward this end and provides evidence for the thesis that the development of phonemic representations and segmental processing at a perceptual level is a protracted process that extends into the beginning school years.

Importantly, the present study also provides empirical support for the hypothesis that the developmental trend from relatively holistic to more segmental processing is a function of vocabulary knowledge and is most appropriately conceptualized as occurring in a word-by-word manner. That is, the increasing need to discriminate from a growing number of similar-sounding alternatives and the increasing frequency of experience with a word propel the developing lexicon toward segmentalized representations. This study is the first developmental investigation of the combined effects of neighborhood structure and frequency on spoken-word recognition performance. Goals for future research include investigating children's perception of varying lexical items across different perceptual tasks, extending the study of neighborhood structures to preschool children, and delineating the relationship between perceptual processing and phonemic awareness via direct and longitudinal comparisons.

REFERENCES

ANDREWS, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18, 234-254.

ASLIN, R. N., PISONI, D. B., & JUSCZYK, P. W. (1983). Auditory development and speech perception in infancy. In M. M. Haith & J. J. Campos (Eds.), *Carmichael's manual of child psychology: Vol. II. Infancy and the biology of development* (pp. 573-687). New York: Wiley.

ASLIN, R. N., & SMITH, L. B. (1988). Perceptual development. *Annual Review of Psychology*, 39, 435-473.

BRADLEY, L., & BRYANT, P. E. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, 271, 746-747.

BRADLEY, L., & BRYANT, P. [E.] (1991). Phonological skills before and after learning to read. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 37-45). Hillsdale, NJ: Erlbaum.

BROWN, G. D. A. (1987). Resolving inconsistency: A computational model of word naming. *Journal of Memory & Language*, 26, 1-23.

BROWN, G. D. A., & WATSON, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory & Cognition*, 15, 208-216.

CARROLL, J. B., & WHITE, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, 25, 85-95.

CHARLES-LUCE, J., & LUCE, P. A. (1990). Similarity neighborhoods of words in young children's lexicon. *Journal of Child Language*, 17, 205-215.

CHARLES-LUCE, J., & LUCE, P. A. (1995). An examination of similarity neighborhoods in young children's receptive vocabularies. *Journal of Child Language*, 22, 727-735.

COLE, R. A. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception & Psychophysics*, 13, 153-156.

COLE, R. A. (1981). Perception of fluent speech by children and adults. In H. Winitz (Ed.), *Native language and foreign language acquisition* (Annals of the New York Academy of Sciences, Vol. 379, pp. 92-109). New York: New York Academy of Sciences

COLE, R. A., & PERFETTI, C. A. (1980). Listening for mispronunciations in a children's story: The use of context by children and adults. *Journal of Verbal Learning & Verbal Behavior*, 19, 297-315.

COTTON, S., & GROSJEAN, F. (1984). The gating paradigm: A compar-

ison of successive and individual presentation formats. *Perception & Psychophysics*, **35**, 41-48.

CUNNINGHAM, A. E. (1990). Explicit versus implicit instruction in phoneme awareness. *Journal of Experimental Child Psychology*, **50**, 429-444.

DOLLAGHAN, C. A. (1994). Children's phonological neighborhoods: Half empty or half full? *Journal of Child Language*, **21**, 257-271.

ELLIOTT, L. L., CLIFTON, L., & SERVI, D. (1983). Word frequency effects for a closed-set word identification task. *Audiology*, **22**, 229-240.

ELLIOTT, L. L., HAMMER, M. A., & EVAN, K. E. (1987). Perception of gated, highly familiar spoken monosyllabic nouns by children, teenagers, and older adults. *Perception & Psychophysics*, **42**, 150-157.

ELLIOTT, L. L., SCHOLL, M. E., GRANT, J. O., & HAMMER, M. A. (1990). Perception of gated, highly familiar spoken monosyllabic nouns by children with and without learning disabilities. *Journal of Learning Disabilities*, **23**, 248-259.

FOWLER, A. E. (1991). How early phonological development might set the stage for phonological awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 97-117). Hillsdale, NJ: Erlbaum.

FOX, R., & KOENIGSKNECHT, R. (1989, November). *The effect of cohort size on gated word recognition.* Paper presented at the meeting of the American Speech and Hearing Association, St. Louis.

GATHERCOLE, S. E., & BADDELEY, A. D. (1993). *Working memory and language.* Hove, U.K.: Erlbaum.

GILHOOLY, K. J. (1984). Word age-of-acquisition and residence time in lexical memory as factors in word naming. *Current Psychological Research & Reviews*, **3**, 24-31.

GILHOOLY, K. J., & GILHOOLY, M. L. (1980). Age-of-acquisition and residence in lexical and episodic memory tasks. *Memory & Cognition*, **7**, 214-223.

GILHOOLY, K. J., & LOGIE, R. H. (1982). Word age-of-acquisition and lexical decision making. *Acta Psychologica*, **50**, 21-34.

GOLDINGER, S. D., LUCE, P. A., & PISONI, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory & Language*, **28**, 501-518.

GROSJEAN, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, **28**, 267-283.

GROSJEAN, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, **38**, 299-310.

HUTTENLOCHER, D. P., & ZUE, V. W. (1983). *Exploring phonotactic and lexical constraints in word recognition.* Paper presented at the 106th meeting of the Acoustical Society of America, San Diego.

JUSCZYK, P. W. (1986). Toward a model of the development of speech perception. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 1-19). Hillsdale, NJ: Erlbaum.

JUSCZYK, P. W. (1992). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17-64). Parkton, MD: York Press.

KUČERA, H., & FRANCIS, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

LANDAUER, T. K., & STREETER, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning & Verbal Behavior*, **12**, 119-131.

LOGAN, J. S. (1992). *A computational analysis of young children's lexicons* (Tech. Rep. No. 8; Research on Spoken Language Processing). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.

LUCE, P. (1986). *Neighborhoods of words in the mental lexicon* (Tech. Rep. No. 6; Research on Speech Perception). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.

LUNDBERG, I. (1991). Phonemic awareness can be developed without reading instruction. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 47-53). Hillsdale, NJ: Erlbaum.

LUNDBERG, I., FROST, J., & PETERSON, O. (1988). Effects of an exten-sive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, **23**, 263-284.

MARSLEN-WILSON, W. D. (1984). Function and process in spoken word recognition: A tutorial review. In H. Bouma & D. G. Bowhuis (Eds.), *Attention and performance: Control of language processes* (pp. 125-150). Hillsdale, NJ: Erlbaum.

MARSLEN-WILSON, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*, **25**, 71-102.

MARSLEN-WILSON, W. D. (1989). Access and integration: Projecting sound onto meaning. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 3-24). Cambridge, MA: MIT Press.

METSALA, J. L. (in press). Spoken word recognition in reading disabled children. *Journal of Educational Psychology*.

NITTROUER, S., & STUDDERT-KENNEDY, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech & Hearing Research*, **30**, 319-329.

NITTROUER, S., STUDDERT-KENNEDY, M., & McGOWAN, R. S. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech & Hearing Research*, **32**, 120-132.

NUSBAUM, H. C., PISONI, D. B., & DAVIS, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words.* (Tech. Rep. No. 10; Research on Speech Perception). Bloomington: Indiana University, Department of Psychology, Speech Research Laboratory.

PISONI, D., & LUCE, P. (1987). Acoustic-phonetic representations in word recognition. *Cognition*, **25**, 21-52.

SALASOO, A., & PISONI, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory & Language*, **24**, 210-231.

SEIDENBERG, M. S., & McCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**, 523-568.

TREIMAN, R., & BREAUX, A. M. (1982). Common phoneme and overall similarity relations among spoken word syllables: Their use by children and adults. *Journal of Psycholinguistic Research*, **11**, 581-610.

TREIMAN, R., & ZUKOWSKI, A. (1991). Levels of phonological awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 67-83). Hillsdale, NJ: Erlbaum.

TYLER, L. K., & WESSELS, J. (1985). Is gating an on-line task? Evidence from naming latency data. *Perception & Psychophysics*, **38**, 217-222.

WAGNER, R. K., TORGESON, J. K., LAUGHON, P., SIMMONS, K., & RASHOTTE, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology*, **85**, 83-103.

WAGNER, R. K., TORGESON, J. K., & RASHOTTE, C. A. (1994). Development of reading related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, **30**, 73-87.

WALLEY, A. C. (1987). Young children's detections of word-initial and -final mispronunciations in constrained and unconstrained contexts. *Cognitive Development*, **2**, 145-167.

WALLEY, A. C. (1988). Spoken word recognition by young children and adults. *Cognitive Development*, **3**, 137-165.

WALLEY, A. C. (1993a). More developmental research is needed. *Journal of Phonetics*, **21**, 171-176.

WALLEY, A. C. (1993b). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, **13**, 286-350.

WALLEY, A. C., & METSALA, J. L. (1990). The growth of lexical constraints on spoken word recognition. *Perception & Psychophysics*, **47**, 267-280.

WALLEY, A. C., & METSALA, J. L. (1992). Young children's age-of-acquisition estimates for spoken words. *Memory & Cognition*, **20**, 171-182.

WALLEY, A. C., MICHELA, V. L., & FLEGE, J. E. (1994, December). *The development of speech perception: Beyond infancy.* Paper presented at the meeting of the Acoustical Society of America, Austin, TX.

WALLEY, A. C., MICHELA, V. L., & WOOD, D. R. (1995). The gating par-

adigm: Effects of presentation format on spoken word recognition by children and adults. *Perception & Psychophysics*, **57**, 343-351.

WALLEY, A. C., SMITH, L. B., & JUSCZYK, P. W. (1986). The role of phonemes and syllables in the perceived similarity of speech sounds for children. *Memory & Cognition*, **14**, 220-229.

NOTES

1. Studies have shown that infants can discriminate among phonemic contrasts (for a review, see Aslin, Pisoni, & Jusczyk, 1983). However, for a discussion of how this early discrimination between individual phonemes differs from later recognition of linguistic units, see Jusczyk's (1992) model (word recognition and phonetic structure acquisition) and Walley's (1993b) developmental review.

2. Logan (1992) has suggested that nonphonemic representation systems may be more economical for young children, while still adequately specifying individual lexical items in their relatively smaller lexicons. The exact form of such developmentally early, more holistic representations remains an unresolved issue in this field. One alternative is that children's early representations may not specify sequential information within a larger syllabic unit. In addition, for young children's smaller lexicons, stress patterns may play a more prominent role in reducing the lexical space from which a target must be discriminated (e.g., Huttenlocher & Zue, 1983). The goal of the present study is not to resolve the issue of what defines more holistic representations beyond previous research. Rather, this paper empirically examines factors hypothesized to be important in the developmental time course of lexical-phonemic restructuring and the trend toward less holistic word-recognition processes.

3. If recognition processes were equivalent across age, it would be predicted that words in children's relatively smaller, more structurally distinct lexicons could be distinguished on the basis of less of the bottom-up input. Indeed, cohort theory (Marslen-Wilson, 1987, 1989) predicts that recognition occurs at that point, on a left-to-right basis, at which a word differs structurally from all other words in the listener's lexicon. Given children's smaller lexicons, predictions from this model would be that children should recognize words at an earlier point than adults.

4. The proportions of blank trials were .064, .046, .039, and .020 (*SD* = .028, .03, .026, .022) for the youngest through adult group, respectively.

5. This operational definition of isolation point is consistent with recent work using the gating paradigm (e.g., Walley et al., 1995). This definition is different from that of Grosjean (e.g., 1985), who first used the term to refer to the gate duration at which the subject correctly guessed the target word without subsequently changing his/her response. Grosjean's definition of isolation point, however, may be more congruent with definitions of the total acceptance point, which has been used to refer to a combination of the gate at which a subject does not subsequently change his/her response and reaches some criterion in a confidence rating of their response (e.g., Elliott et al., 1987; Walley et al., 1995).

6. This addition of 50 msec, the duration of one gate, for calculating an isolation point when a target was not correctly identified has been used in previous research by Elliott et al. (1987) and Walley et al. (1995). Even adults sometimes need additional information to identify words in isolation (e.g., the onset of the next word; Grosjean, 1985). This method for nonrecognition correction is liberal, as it assumes the subject would have recognized the word given an offset cue. Indeed,

analogous analysis conducted with the elimination of nonrecognition trials did not change the pattern of results (see note 7).

7. The same analysis was completed with mean recognition times calculated over only those trials for which the subject correctly identified the target word. In other words, trials for which the subject never guessed the target word correctly (i.e., errors) were eliminated. The pattern of results was the same as for the main analysis, with the exception that the three-way interaction did not reach traditional levels of statistical significance ($F = 2.39$, $p < .08$). Age did interact with frequency ($F = 3.16$, $p < .05$) and with density ($F = 4.77$, $p < .005$), and the conclusions from these interactions are consistent with the main analysis reported in the body of the paper. Developmental differences were least pronounced for words in dense neighborhoods.

APPENDIX
Target Words and Neighborhood Properties

Word	Word Type	Log Frequency	Number of Neighbors	Mean Log Frequency of Neighbors
cut	high–dense	3.28	16	1.81
sign	high–dense	2.97	11	2.61
ship	high–dense	2.92	16	1.81
bill	high–dense	3.16	13	2.53
head	high–dense	3.63	14	2.47
case	high–dense	3.56	14	2.49
bag	high–dense	2.62	20	1.83
girl	high–sparse	3.34	5	1.51
dirt	high–sparse	2.63	6	2.55
church	high–sparse	3.54	4	1.45
voice	high–sparse	3.35	2	2.75
bar	high–sparse	2.91	5	2.79
dog	high–sparse	2.87	6	2.09
sky	high–sparse	2.76	3	1.78
den	low–dense	1.30	14	2.79
comb	low–dense	1.78	12	2.56
chum	low–dense	1.00	12	1.82
lace	low–dense	1.85	14	2.76
mug	low–dense	1.00	14	1.65
weed	low–dense	1.00	11	1.98
vet	low–dense	1.00	11	2.83
toad	low–sparse	1.60	2	2.62
chime	low–sparse	1.00	3	1.57
claw	low–sparse	1.00	3	1.82
stew	low–sparse	1.70	2	2.45
mole	low–sparse	1.60	4	2.07
fudge	low–sparse	1.00	5	1.98
leash	low–sparse	1.48	4	2.52

Note—Word types *high* and *low* refer to words of high- versus low-frequency stimulus cells; *dense* and *sparse* refer to neighborhood structure: words from dense neighborhoods have many similar-sounding words in the listener's lexicon as opposed to sparse words, which have few similar-sounding words in the listener's lexicon. Frequency counts based on Kučera and Francis (1967).