

# Three- and four-year-olds' perceptual confusions for spoken words

LOUANN GERKEN

*State University of New York, Buffalo, New York*

WALTER D. MURPHY

*Vanderbilt University, Nashville, Tennessee*

and

RICHARD N. ASLIN

*University of Rochester, Rochester, New York*

Although infants have the ability to discriminate a variety of speech contrasts, young children cannot always use this ability in the service of spoken-word recognition. The research reported here asked whether the reason young children sometimes fail to discriminate minimal word pairs is that they are less efficient at word recognition than adults, or whether it is that they employ different lexical representations. In particular, the research evaluated the proposal that young children's lexical representations are more "holistic" than those of adults, and are based on overall acoustic-phonetic properties, as opposed to phonetic segments. Three- and four-year-olds were exposed initially to an invariant target word and were subsequently asked to determine whether a series of auditory stimuli matched or did not match the target. The critical test stimuli were nonwords that varied in their degree of phonetic featural overlap with the target, as well as in terms of the position(s) within the stimuli at which they differed from the target, and whether they differed from the target on one or two segments. Data from four experiments demonstrated that the frequency with which children mistook a nonword stimulus for the target was influenced by extent of featural overlap, but not by word position. The data also showed that, contrary to the predictions of the holistic hypothesis, stimuli differing from the target by two features on a single segment were confused with the target more often than were stimuli differing by a single feature on each of two segments. This finding suggests that children use both phonetic features and segments in accessing their mental lexicons, and that they are therefore much more similar to adults than is suggested by the holistic hypothesis.

Since the pioneering work of Eimas, Siqueland, Jusczyk, and Vigorito (1971), it has been demonstrated that infants as young as two months of age have sufficient perceptual abilities to distinguish a variety of speech contrasts (see Aslin, Pisoni, & Jusczyk, 1983; Best, McRoberts, & Sithole, 1988; Eimas, Miller, & Jusczyk, 1987; Jusczyk, 1985; Kuhl, 1987; Werker, 1991). In contrast, studies with young children suggest that they cannot always apply these perceptual abilities directly to the problem of spoken-word recognition (Barton, 1976a, 1976b, 1980; Edwards, 1974; Eilers & Oller, 1976; Garnica, 1971; Graham & House, 1971; Shvachkin, 1948/1973). In one study, Eilers and Oller (1976) taught 22- to 26-month-olds to refer to a toy with a nonsense word (*tig*) that created a minimal pair with a familiar word (*pig*).

They found that, on average, children made perceptual discrimination errors 36% of the time, in a situation in which adults would presumably make no errors. Thus, although the literature on infant speech perception suggests that children have the auditory capabilities to perceptually distinguish such contrasts as /p/ versus /t/, they nevertheless perform much worse than adults in word discrimination. What is the nature of children's relatively poor ability to discriminate among spoken words?

Perhaps young children fail to consistently perceive phonetic differences that they discriminated in infancy because the tasks used to test discrimination ability in children are more demanding than the tasks used with infants (e.g., Jusczyk, 1992; Locke, 1988). For infants to be credited with discrimination of a particular speech contrast, they need only distinguish a new stimulus from an old one; in contrast, speech-perception experiments with young children typically require them to match an acoustic string with its referent (but see Graham & House, 1971), and children's performance on these tasks therefore reflects not only their perceptual abilities, but also their ability to recall the referent of an auditory stimulus. The fact that children perform significantly better

---

This research was supported by National Science Foundation Grant DBS91020952 to L.A.G. and by National Institutes of Health Grant HD20286 to R.N.A. We thank Mary Boyle, Robin Cooper, Jill Gallipeau, and two anonymous reviewers for helpful comments and assistance. Correspondence should be addressed to L. A. Gerken, Department of Psychology, Park Hall, SUNY at Buffalo, Buffalo, NY 14260 (e-mail: psygerkn@ubvms.cc.buffalo.edu).

on discrimination tasks when both words of a test pair are in their expressive vocabulary suggests that general cognitive demands of the referent identification task contribute to poor performance (Barton, 1976a). A similar explanation might account, in part, for differences in word-discrimination performance between children and adults. Children's performance might be poorer because of developmental differences in memory, attention, and information-processing capacity, independent of speech-perception abilities per se. Such an explanation for child/adult differences leaves open the possibility that both groups of listeners employ the same perceptual representations in spoken-word recognition.

Alternatively, however, children's relatively poor performance in word-recognition tasks might reflect their use of *different* lexical representations from those used by adults. Several researchers have suggested that when they are learning language, children's primary goal is to recognize and produce whole words, not to learn phonemic contrasts per se (e.g., Chiat, 1979; Fey & Gandour, 1982; Jusczyk, 1992). These researchers have further suggested that, in keeping with this goal, children represent early words in terms of *holistic* properties, such as prosodic structure and acoustic shape, or in terms of phonetic features that are not bundled into individual segments (e.g., Charles-Luce & Luce, 1990; Ferguson & Farwell, 1975; Jusczyk, 1982, 1986; Logan, 1992; Macken, 1979; Menyuk & Menn, 1979; Studdert-Kennedy, 1986; Treiman & Baron, 1981; Treiman & Breaux, 1982; Vihman & Velleman, 1989; Walley, 1993; Waterson, 1971). According to this view, children discriminate spoken words on the basis not of particular phonemic contrasts, but of overall acoustic or phonetic featural differences.

While some researchers have suggested that lexical representations are holistic only until the vocabulary growth spurt begins, at about two years of age (Ferguson, 1986; Menyuk & Menn, 1979; Studdert-Kennedy, 1987), others have proposed a *developmentally protracted holistic hypothesis*—namely, that children's lexical representations are different from those of adults until the early school years (Treiman & Baron, 1981; Treiman & Breaux, 1982; Walley, 1987, 1988, 1993; Walley, Smith, & Jusczyk, 1986). It is this hypothesis that will be addressed in our studies. According to this account, children decrease their use of holistic word representations (either acoustic/prosodic word shapes or phonetic features not bundled into segments) in favor of segmental representations as a result of one or more of the following developmental changes: maturation; exposure to segmental representations in learning to read (Treiman & Baron, 1981; Treiman & Breaux, 1982; Walley, 1993); increased familiarity with, or frequency of exposure to, individual words (Walley, 1993; Walley & Metsala, 1990); or acquisition of pairs or sets of lexical items that differ by only a single phonetic segment (Charles-Luce & Luce, 1990; Walley, 1993; also see Dollaghan, 1994).

One form of evidence that has been taken to support the developmentally protracted holistic hypothesis is the

finding that children are less able than adults to employ segmental representations of words in similarity judgments or phonetic manipulation tasks (Stanovich, Cunningham, & Cramer, 1984; Treiman & Breaux, 1982). Further evidence is that children show greater use of coarticulation than adults in both production and perception, indicating that their lexical representations may comprise larger units than those of adults (Nittrouer & Studdert-Kennedy, 1987; Nittrouer, Studdert-Kennedy, & McGowan, 1989). Still more evidence comes from a task in which noise either replaces a segment in a word or is added to it. Walley (1988) asked five-year-olds and adults to listen to such stimuli and found that adults, but not children, rated stimuli as "noisier" when noise replaced the initial segment. According to the holistic hypothesis, adults find word-initial noise more disruptive because they accord a special status to word-initial segments in their lexical representations. In contrast, children fail to accord a special status to word-initial segments, either because their early word representations are not based on segments at all, or because they do not employ segmental representations during lexical access. Yet another source of evidence supporting the developmentally protracted holistic hypothesis comes from gating studies in which listeners are presented, across trials, with larger and larger acoustic "slices" of a word (Grosjean, 1980). Research using this technique has demonstrated that five- to six-year-old children require significantly more acoustic information than adults to identify even highly familiar words (Elliott, Hammer, & Evan, 1987; Walley, 1988). This finding has been taken to suggest that, in children's lexical representations, crucial distinguishing information is distributed across the whole word, while in adults' representations, this information is more localized to the beginning segment or segments (also see Treiman & Baron, 1981; Treiman & Breaux, 1982).

There are several problems with the developmentally protracted holistic hypothesis. The first concerns children's ability to explicitly access word segments for purposes of similarity judgments or other metalinguistic tasks. Although this ability is almost certainly an important underpinning of literacy (Stanovich et al., 1984; Treiman & Breaux, 1982), the lack of such an ability by no means implies the lack of a segmental representation (e.g., Walley, 1993).

The second problem with the developmentally protracted holistic hypothesis concerns children's early language productions, which suggest that they employ segmental representations quite similar to those posited for adults. In particular, like those of adults, children's slips of the tongue involve whole segments (Gerken, in press; Jaeger, 1992; Stemmer, 1989; Wijnen, 1992). For example, a child who attempts to say "big dog" and makes the slip "dig dog" has anticipated the segment /d/ and substituted it for the segment /b/. Such slips have been found in children as young as 17 months (Jaeger, 1992) and indicate that although children may coarticulate an initial consonant with a following vowel more than adults

do (Nittrouer et al., 1989), the consonant and vowel are represented separately, and not as parts of diphones or other larger-than-segment units. However, the lexical representations or processes that children use to produce speech may be different in important ways from the ones they initially use for spoken-word recognition (Anisfeld, 1984; Gerken, 1994; Menn, 1978, 1980, 1983; Spencer, 1986; Straight, 1980; Vihman, 1993). Also, by definition, children make slips on words that they know at least well enough to produce, and word familiarity has been proposed as a factor that encourages the use of segmental representations (Walley, 1993; Walley & Metsala, 1990). Therefore, by themselves, the slip data do not necessarily rule out a holistic account of early word recognition.

Another problem with the developmentally protracted holistic hypothesis is that the primary supporting evidence involves differences between adults' and children's attention to word-initial segments (Treiman & Baron, 1981; Treiman & Breaux, 1982; Walley, 1987, 1988; Walley et al., 1986). These differences are important in the context of theories of adult word recognition, whereby candidate words are proposed and eliminated as information arrives in a "left-to-right" fashion (e.g., Marslen-Wilson, 1989; Marslen-Wilson & Welsh, 1978). According to this view, the first one or two segments of a word are crucial for identifying a cohort of candidate words from which the correct word is eventually chosen. Children's failure to accord special status to word-initial segments has therefore been taken to suggest that they do not identify words in a left-to-right, segment-by-segment fashion.

However, other recent accounts do not view adult spoken-word recognition as a strictly left-to-right process in which initial segments necessarily play the primary role (e.g., Luce, 1986; Luce, Pisoni, & Goldinger, 1990). Rather, these accounts construe the listener's job as one of distinguishing a word from phonetically similar neighbors. Within the neighborhood framework, the question becomes one of which acoustic/phonetic properties make a word more or less distinguishable from all others, and not segment position in the word per se (e.g., Goldinger, Luce, & Pisoni, 1989; Luce, Goldinger, & Auer, in press). Evidence favoring the neighborhood approach over left-to-right accounts comes from the fact that not all listening tasks with adults demonstrate a special status for word-initial segments (Grosjean, 1985; Luce, 1986; Luce et al., 1990). For example, adults who are presented with two words and asked to determine whether they are the same or different do not accord special status to word-initial segments in their decision, but are influenced by overall featural similarity (Goldinger et al., 1989; Luce et al., in press). These varied findings raise the possibility that in some experimental tasks adults may adopt strategies that emphasize word-initial segments, whereas in others they may not. Thus, apparent differences between children's and adults' attention to word-initial segments may not reflect differences in their use of segmental representations, but, rather, they may reflect develop-

mental differences in the extent to which the two groups spontaneously adopt such strategies on a given task.

In summary, the most clear and robust differences between early childhood and adulthood, in terms of spoken-word recognition, are that over development there is, on the one hand, a general improvement in the ability to accurately distinguish minimal word pairs, and on the other hand, a greater reliance on the information carried by word-initial segments for certain tasks. We have suggested that the former difference may reflect developmental changes in general cognitive abilities demanded by the referent identification task, and not changes in spoken-word representation per se, and similarly, that the latter child/adult difference may reflect differential use of task-specific strategies (e.g., Cole & Perfetti, 1980).

To resolve some of the problems with previous studies of spoken-word recognition in children, and to provide further information about their lexical representations, we need a task that does not require children to associate words with referents and that does not differentially encourage the use of task-specific strategies by children and adults. A two-stimulus comparison task, such as the one used by Luce (1986; Luce et al., 1990), might meet these criteria, because it has no explicit referential component, and because adults appear to base their responses on overall similarity and do not seem to accord special status to word-initial segments. However, this task requires listeners to remember the first stimulus and compare it with the second on each trial, and therefore demands considerable attention and memory. Furthermore, a listener could perform the task simply by gauging the *physical* similarity of the two stimuli, and not necessarily by using lexical representations at all.

Therefore, all of our experiments employed a modified version of the two-stimulus comparison task, in which three- and four-year-old children were asked to distinguish an *invariant* target word from phonetically similar nonword foils (i.e., an AX-like task). Children were required to remember only a single word throughout the experiment in order to reduce attention and memory demands in comparison with the requirements of the normal two-stimulus comparison procedure. Additionally, because children always compare a stimulus to a mental representation of the target, the task was intended to discourage comparisons of the target and foils that were based on physical similarity (since in this task, the target is not physically present), and to increase the likelihood that children's performance will reflect their lexical representations. This task allowed us to ask which phonetic properties of a foil would make it more or less confusable with the target. Experiment 1 explored whether the properties of the foils that influence adults' responses are different from those that influence children's responses; the results suggest that both groups of listeners are influenced by the featural overlap of target and foils. Experiment 2 further confirmed the influence of featural overlap on children's performance by ruling out several alternative accounts. Experiment 3

employed a new stimulus set to test the generalizability of the featural-overlap hypothesis and to further examine the discriminability of consonant features versus vowel features. Finally, Experiment 4 asked whether children's lexical representations consist of phonetic features that are bundled into segments. The results from these studies indicate that although children's spoken-word recognition may be less efficient than that of adults, their lexical representations are closer to those of adults than is suggested by the holistic account of early word recognition.

### EXPERIMENT 1

In Experiment 1, the target was the word *little*, and the foils differed from the target on either the first, the second, or the third segment (i.e., the onset, the vowel, or the coda, respectively). Thus, the foils were /nItI/, /IEtI/, and /IIgI/. We reasoned that a listener might approach our task in at least two ways, thereby yielding two distinct patterns of reaction-time and false-alarm data. One such approach might be to maintain a memory representation of the target word in terms of segments and, as a stimulus is presented, determine in a left-to-right fashion whether each incoming segment matches the one in memory. The segment-matching approach is consistent with proposals in which initial segments have a special status in lexical access (Marslen-Wilson & Welsh, 1978; Walley, 1993). It would lead to rejections of stimuli differing from the target on early segments being faster than rejections of those differing on later segments, because early segments appear earlier in time. Thus, the pattern of reaction times for correctly judging that the foils are different from the target should be: /nItI/ < /IEtI/ < /IIgI/. The segment-matching approach might also lead a listener to make a target response incorrectly (false alarm) when one or more segments from the beginning of the foil matched the target; foils differing from the target on a later segment might therefore elicit more target responses (false alarms) than foils differing on an earlier segment, and the pattern of false alarms should be the same as the pattern of reaction times.

A second approach to the task might be to maintain a memory representation of the target in terms of phonetic features and compare the overall featural similarity of the stimulus on each trial with that of the target. The featural-overlap approach is consistent with data from the two-stimulus comparison task used by Luce (1986; Luce et al., 1990). It would lead to rejections of stimuli differing from the target on a greater number of features being faster than rejections of those differing on fewer features. Consider the featural similarity of the target *little* and the three foils (see Table 1). The /n/ in /nItI/ differs from the initial /l/ in *little* by only a single feature. In contrast, the /t/ in *little* and the /g/ in /IIgI/ differ by two features. The /l/ in *little* and the /E/ in /IEtI/ also differ by only a single feature; however, there is no empirical evidence that vowel and consonant features can be equated in adult listeners.<sup>1</sup> If we consider only consonant features, the pattern of reaction times should be

**Table 1**  
Featural and Segmental Relations of Targets and Foils in Experiments 1–4

Target	Foil				
	Differences from Target				
	Features		Segments		
	Number	Type	Number	Type	
Experiments 1 and 2					
<i>little</i>	/nItI/	1	Manner	1	Onset
	/IEtI/	1	Vowel height	1	Vowel
	/IIgI/	2	Place, voice	1	Coda
	/gItI/	2	Place, manner	1	Onset
	/ItIn/	2	Manner, glottal stop	1	Coda
Experiments 3 and 4					
<i>lick</i>	/nIk/	1	Manner	1	Onset
	/IEk/	1	Vowel height	1	Vowel
	/IIg/	1	Voice	1	Coda
	/gIk/*	2	Place, manner	1	Onset
	/IIf/*	2	Place, manner	1	Coda
	/zIg/*	2	Manner of onset, voice of coda	2	Onset, coda
/zEk/*	2	Manner of onset, vowel height	2	Onset, vowel	

\*Crucial test cases for the features-and-segments hypothesis (see text).

/IIgI/ < /nItI/. The featural-overlap approach would also lead to more false alarms for foils that share more features with the target, and the pattern of false alarms should be the same as the pattern of reaction times. Note that both the reaction-time and the false-alarm predictions of the featural-overlap hypothesis are in the opposite direction to the predictions of the segment-matching (left-to-right) hypothesis.

On the basis of these two hypotheses about how listeners approach our task, the strongest evidence *favoring* the holistic hypothesis would be provided by adults producing the pattern of data predicted by the segment-matching hypothesis and by children producing the pattern predicted by the featural-overlap hypothesis. Such an outcome would indicate that adults employ segmental representations, while children do not. In contrast, the strongest evidence *against* the holistic hypothesis would be if both adults and children were to produce the pattern of data predicted by the segment-matching hypothesis. This outcome would indicate that both groups of listeners employ segmental representations. A third outcome is also possible—namely, one in which both adults and children produce the pattern of data predicted by the featural-overlap hypothesis. In order to assess the implications of such an outcome for the holistic hypothesis, further experiments would be needed to determine whether the similarity in the two groups' performance reflected similar lexical representations.

### Method

**Subjects.** The subjects for all experiments were recruited from birth announcements in Rochester newspapers and were reported by their parents to have normal hearing. Different groups of sub-

jects were tested in the four experiments. Seventeen 3- to 4-year-old children (10 males and 7 females, age range = 47–56 months, mean = 50) participated in Experiment 1. An additional seven children were tested but not included in the study because they failed to meet the inclusion criterion (see below). Nine college students also participated to provide an adult comparison group.

**Materials.** The target word *little* was chosen because it is familiar to children and likely to be easily remembered in our task. Also, when it is produced in its citation form, *little* has no minimal pairs in English; therefore, according to the view that children come to use segmental representations on a piecemeal basis as they acquire pairs or sets of phonetically similar words (e.g., Charles-Luce & Luce, 1990; Walley, 1993), *little* should still be represented holistically by children of the ages we tested.<sup>2</sup> As noted in the Introduction, three nonword foils to *little* were created by changing either the onset, the vowel, or the coda of the first syllable of this word, yielding /nltl/, /lEt/, and /lgl/, respectively. Nonword foils were used (as opposed to words) in order to avoid having sets of stimuli that created minimal word pairs and that might thereby have increased the likelihood that children would have a segmental representation of the words in the set (see above). In addition to the target and foils, four other English words and two nonwords, each of which was phonetically dissimilar from either the target word or the foils, were included as control stimuli. These were *cookie*, *teacher*, *water*, *Mickey*, /gaeldn/, and /garbn/. The target word, foils, and control stimuli in this and subsequent experiments were recorded by a female talker and digitized at 10 kHz with a 4.8-kHz lowpass filter for computer presentation.

**Procedure.** Children were brought by their parents to the Developmental Speech Perception Laboratory at the University of Rochester. They were taken into an IAC sound-attenuated booth containing a child-sized desk, a loud speaker placed directly in front of the desk, and a “talking” dog. Parents were allowed to sit in a chair three feet behind the child in the booth, and were asked to refrain from speaking during the experiment. The desk top contained two pressure-sensitive plates that could detect a “press” from the child’s hand. This allowed us to collect children’s responses, as well as their response latency, automatically. On either side of the desk top were stickers, one showing a smiling face and the other showing a frowning face; the positions were counter-balanced across subjects. The children were told that the dog was going to say some “things,” and that sometimes it would say the word *little*. If it did, the child should press the smiling face as fast as he or she could, but if the dog said anything else, the child was to press the frowning face.

The children were then taken through ten practice trials in which the stimuli consisted of the target word and the control stimuli. None of the three foils was presented during the practice trials. Children who did not correctly respond to at least 90% of the practice stimuli were not included in the study; those who met the criterion participated in the three blocks of test trials. Each 12-trial block consisted of 3 trials of the target, interspersed among 3 trials using the foils (i.e., 1 trial for each of the foils) and 6 trials using the control stimuli (i.e., 1 trial for each of the control stimuli). The stimuli were presented at 70 dB (SPL), both in practice and in experimental trials. Each trial consisted of a single presentation of the word or nonword. The same procedure was followed in all four experiments, although the number of trials per block varied between experiments, depending on the number of foils and control stimuli used.

## Results and Discussion

Two dependent variables were examined—namely, reaction times for correct responses (see Table 2) and percentage of target responses (see Table 3). For the reaction-time measure, a *t* test demonstrated that adults responded significantly faster than children [1,056 msec

**Table 2**  
Reaction Times (in Milliseconds) for Correct Responses  
in Experiments 1 and 3

Stimulus	Reaction Time			
	Children		Adults	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 1				
<i>little</i>	2,250	455	944	160
/nltl/	2,844	583	1,154	493
/lEt/	2,543	750	1,127	207
/lgl/	2,607	1,104	1,092	169
Controls	2,231	324	962	178
<i>M</i>	2,495	726	1,056	274
Experiment 3				
<i>lick</i>	1,674	288	953	189
/nltk/	2,146	563	1,027	142
/gltk/	2,071	432	977	121
/lEk/	2,504	690	1,332	367
/llf/	2,186	537	1,063	131
Controls	2,346	781	977	145
<i>M</i>	2,154	612	1,055	228

Note—*M*, mean; *SD*, standard deviation.

vs. 2,495 msec,  $t(24) = 5.69$ ,  $p < .0001$ ]. Unfortunately, it was not possible to perform an analysis comparing children’s reaction times with the different types of stimuli. This was because there was a large number of missing data points (26%), largely due to the failure of 11 children to produce any correct responses for one or more of the foils.

A one-way analysis of variance (ANOVA) was performed on the adults’ reaction-time data, comparing their responses to the target word, to the three foils, and to the control stimuli. The ANOVA showed a significant effect of stimulus type [ $F(4,32) = 3.20$ ,  $p < .05$ ]. Pairwise comparisons (Newman-Keuls,  $p = .05$ ) showed that

**Table 3**  
Percentage of Target Responses in Experiments 1 and 2

Stimulus	Target Responses (%)			
	Children		Adults*	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 1				
<i>little</i>	91	15	100	0
/nltl/	65	45	8	15
/lEt/	45	42	0	0
/lgl/	29	44	4	12
Controls	6	9	0	0
<i>M</i>	47	45	23	40
Experiment 2				
<i>little</i>	90	13		
/nltl/	47	37		
/lEt/	27	34		
/lgl/	24	29		
/gltl/	29	38		
/lltn/	24	29		
Controls	3	4		
<i>M</i>	35	38		

Note—*M*, mean; *SD*, standard deviation. \*Adult subjects were not used in Experiment 2.

adults responded more slowly to the foils than they did to the target. Not surprisingly, they also responded to the three foils more slowly than they did to the phonetically dissimilar control words. The latter result suggests that adults found the target to be more confusable with the foils than with the control stimuli. Although adults did not demonstrate significant differences in their reaction times to the three foils, it is important to note that the pattern of times ( $/lgl/ < /nltl/$ ) is the opposite of that predicted by the segment-matching hypothesis, whereas it is exactly the one predicted if listeners respond on the basis of the featural overlap of the target and foils.<sup>3</sup>

For the percentage of target responses, a two-way ANOVA, with age (child vs. adult) as a between-subjects factor and stimulus (target vs. foil vs. control) as a within-subjects factor, showed that, across all stimuli, adults made fewer target responses than children [23% vs. 47%,  $F(1,24) = 8.96$ ,  $p < .01$ ]. The analysis also showed a significant effect of stimulus [ $F(4,96) = 45.48$ ,  $p < .0001$ ] and a significant age  $\times$  stimulus interaction [ $F(4,96) = 6.42$ ,  $p < .001$ ]. Not surprisingly, pairwise comparisons (Newman-Keuls,  $p = .05$ ) of the adult data showed that they made more target responses to the actual target word than to the other stimuli; there was no difference, however, between their responses to the foils and their responses to the control stimuli.

Pairwise comparisons of children's target-response data provide a more telling picture of their target-foil confusions. Like adults, children made more target responses to the target word than to the other stimuli. In addition, however, children made more target responses to the foils than to the control stimuli, and among the foils, they made more target responses to  $/nltl/$  than to the other two foils ( $/letl/$  and  $/lgl/$ ). Thus, the pattern of children's target responses is the one predicted if listeners respond on the basis of the number of overlapping consonant features between the target and the foils.

Consistent with other developmental research on the relation between speed and accuracy in adults and children (e.g., Kail, 1991), the following two aspects of children's target responses correspond to the reaction times of the adults: (1) children made more target responses (false alarms), and adults made slower responses, to foils than to control stimuli, which suggests that both groups of listeners were affected by the phonetic similarity of the stimuli and found it more difficult to distinguish foils than to distinguish control stimuli from the target; and (2) the children's pattern of responses to the three foils was similar to the adults'; in particular, children reliably produced more target responses to the foil  $/nltl/$  than to the other two foils. Although there were no significant differences among adults' reaction times to the three foils, their pattern of responding was similar to the children's pattern of false alarms, with  $/nltl/$  receiving longer responses than the other foils. Thus, the pattern of responses across age is most consistent with the featural-overlap account. Before we can assess the implications for the holistic hypothesis, we must determine

whether the similar response patterns of adults and children reflect the use of similar lexical representations.

There are several possible explanations of the children's target-foil confusions, and in particular of why they made more target responses to  $/nltl/$  than to the other foils. One explanation already discussed is that  $/nltl/$  shares more consonant features with the target than do the other foils; the fact that  $/nltl/$ , which differed from the target by a single consonant feature, received more target responses than  $/letl/$ , which differed by a single vowel feature, might be taken to indicate that consonants differing by a single feature are more confusable than vowels differing by a single feature. Another explanation concerns the number of segments shared by the target and individual foils: The foil  $/nltl/$  contains the longest uninterrupted string of segments ( $/ltl/$ ) matching the target. Finally, there is the explanation that  $/nltl/$  shares the rime (vowel and coda) of the first syllable of the target.

Of these three possible accounts—featural overlap, longest uninterrupted string of segments, and matching rime—the first is closest to representational models that have been proposed for adult listeners. For example, Miller and Nicely (1955) showed that the degree of featural overlap between two phonetic segments largely determined how perceptually confusable they were. More recent research on adult spoken-word recognition also supports the important role of features (e.g., Cole, 1973, 1981; Luce et al., in press; Milberg, Blumstein, & Dworetzky, 1988). Therefore, if children's performance on our task proves to have been based on featural overlap between the foils and the target, we would have evidence that adults and children employ similar lexical representations. Importantly, we would also have evidence *against* one form of the developmentally protracted holistic hypothesis, which proposes that children's lexical representations are based on overall acoustic/prosodic (as opposed to featural) properties of words.<sup>4</sup> Experiment 2 was therefore designed to determine whether the featural-overlap hypothesis provides the best account of children's performance in Experiment 1.

## EXPERIMENT 2

This experiment was designed to contrast the featural-overlap hypothesis, the longest-uninterrupted-string-of-segments hypothesis, and the matching-rime hypothesis. To accomplish this, we included the same three foils that were used in Experiment 1 ( $/nltl/$ ,  $/letl/$ , and  $/lgl/$ ), as well as two new foils,  $/gltl/$  and  $/ltn/$ . The featural-overlap hypothesis predicts that children should make the greatest number of incorrect target responses (false alarms) to  $/nltl/$ . Crucially, they should make more target responses to  $/nltl/$  than to  $/gltl/$ , because  $/n/$  differs from  $/l/$  by a single feature, whereas  $/g/$  differs from  $/l/$  by two features (see Table 1). In contrast, the longest-uninterrupted-string-of-matching-segments hypothesis predicts that children should make the greatest number of target responses (false alarms) to  $/nltl/$ ,  $/gltl/$ , and  $/ltn/$ , because

all three foils share an uninterrupted string of three segments with *little*.<sup>5</sup> Finally, the matching-rime hypothesis predicts that children should make the greatest number of target responses (false alarms) to both /nItl/ and /gItl/, because both share the target's rime.

### Method

**Subjects.** Fifteen 3- to 4-year-old children (10 males and 5 females, age range = 49–53 months, mean = 51) participated in Experiment 2. Four additional children were tested, but failed to meet the inclusion criterion (see Experiment 1).

**Materials.** The control stimuli in Experiment 2 were *teacher*, *monkey*, /hodn/, and /sæki/ (some of the control stimuli were changed from those used in the previous experiment to avoid phonetic similarity with the new foils, /gItl/ and /Ittn/).

### Results and Discussion

Because the children's reaction times in Experiment 1 were relatively uninformative, in Experiment 2, we focused on the percentage of target responses (see Table 3). A one-way ANOVA, with stimulus (target vs. foil vs. control) as a within-subjects variable, revealed a significant effect of stimulus [ $F(6,84) = 19.46, p < .0001$ ]. Planned, one-tailed *t* tests were used to test the predictions of the three hypotheses under investigation. Consistent with the featural-overlap hypothesis, children made more target responses to the foil /nItl/ (47%) than to /gItl/ [29%,  $t(14) = 2.00, p < .05$ ], /Ilg/ [24%;  $t(14) = 2.67, p < .01$ ], or /Ittn/ [24%,  $t(14) = 2.67, p < .01$ ]. They also made more target responses to /nItl/ than to /IEtI/ [27%,  $t(14) = 2.33, p < .025$ ]. The fact that the false-alarm rates were larger for /nItl/ than for /gItl/ is consistent with neither the longest-uninterrupted-string-of-segments hypothesis nor the matching-rime hypothesis.

The data from Experiments 1 and 2 suggest that children represent spoken words in terms of phonetic features, just as has been proposed for adults (e.g., Goldinger et al., 1989; Luce et al., in press; Milberg et al., 1988). Our results are similar to those of Graham and House (1971), who asked three- and four-year-olds to make *same-different* judgments for pairs of three-syllable nonsense words. They found that children were most likely to make incorrect "same" judgments when the two stimuli differed by only a single feature. However, our findings further indicate that phonetic features are important in predicting children's discrimination ability when one of the items in a minimal pair is a real word that is not presented on each trial and must be represented in memory. This situation is more similar to the one found in actual lexical access. Furthermore, our task made it unlikely that children's performance was based only on the physical similarity of the stimuli, and thereby increases the likelihood that we were tapping children's lexical representations. The fact that both adults and children appear to be influenced by featural overlap in their discrimination performance suggests that they may employ approaches to spoken-word recognition that are more similar than has been proposed by some supporters of the holistic hypothesis.

The data from Experiments 1 and 2 raise two questions about the generalizability of the featural-overlap account. The first concerns whether children are also influenced by the featural overlap of target and foils when faced with monosyllabic stimuli, or with a target that forms many minimal pairs with real English words. The second question concerns the status of vowel features in children's and adults' lexical representations; recall that the foil /IEtI/ differed from the target by a single vowel feature, but that it nevertheless received fewer target responses than did /nItl/, which differed by one consonant feature. Was the relative discriminability of consonant and vowel features that was observed in these experiments specific to our stimuli? These questions were addressed in Experiment 3.

### EXPERIMENT 3

Experiment 3 employed only monosyllabic stimuli. The target word was *lick*, and foils differed from the target by either a single consonant feature (/nik/),<sup>6</sup> a single vowel feature (/IEk/), or two consonant features (/gIk/ and /IIf/; see Table 1). Unlike the target used in Experiments 1 and 2, the new target forms minimal pairs with many words that young children are likely to know (e.g., "kick," "lock," "lip," etc.). The first question addressed in Experiment 3 was therefore whether, with these new stimuli, children and adults would continue to show evidence in support of the featural-overlap hypothesis. This hypothesis predicts that listeners should respond more slowly and make more target responses (false alarms) to /nik/ than to either /gIk/ or /IIf/. This is because /n/ and /l/ differ by only a single phonetic feature, while /g/ differs from /l/ and /f/ differs from /k/ by two features. The second question addressed in Experiment 3 concerned children's and adults' responses to /IEk/: Would they produce more target responses to foils that differ from the target by a single consonant feature (/nik/) than to those that differ by a single vowel feature (/IEk/), as they did in Experiments 1 and 2?

### Method

**Subjects.** The subjects were fourteen 3- to 4-year-old children (8 males and 6 females, age range = 49–54 months, mean = 51). Six adults also participated as control subjects.

**Materials.** In addition to the target and foils noted above, two words (*teach* and *car*) and two nonword controls (/hod/ and /faem/) were used.

### Results and Discussion

As in Experiment 1, we examined adults' and children's reaction times for correct responses (see Table 2) and examined their percentage of target responses (see Table 4). Turning first to the reaction times, a two-way ANOVA, with age (child vs. adult) as a between-subjects factor and stimulus (target vs. foils vs. control) as a within-subjects factor, revealed a significant effect of age [ $F(1,18) = 33.94, p < .00001$ ], with adults responding more quickly than children (1,055 msec vs. 2,154 msec).

**Table 4**  
**Percentage of Target Responses in Experiments 3 and 4**

Stimulus	Target Responses (%)			
	Children		Adults*	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 3				
<i>lick</i>	99	3	100	0
/nIk/	38	41	0	0
/IEk/	38	41	0	0
/gIk/	21	31	0	0
/IIf/	24	40	0	0
Controls	5	7	0	0
<i>M</i>	38	43	17	38
Experiment 4				
<i>lick</i>	98	2		
/nIk/	44	47		
/IEk/	56	47		
/Ilg/	51	45		
/gIk/	49	43		
/IIf/	40	40		
/zIg/	29	40		
/zEk/	13	30		
Controls	3	4		
<i>M</i>	43	44		

Note—*M*, mean; *SD*, standard deviation. \*Adult subjects were not used in Experiment 4.

There was also a significant effect of stimulus [ $F(5,90) = 5.55, p < .001$ ]. Planned *t* tests were used to address two specific questions: (1) With respect to whether foils differing from the target by one consonant feature elicited slower reaction times than those differing by two consonant features, *t* tests revealed no significant differences between /nIk/ (children = 2,146 msec, adults = 1,027 msec) and either /gIk/ (children = 2,071 msec, adults = 977 msec) or /IIf/ (children = 2,186 msec, adults = 1,063 msec); and (2) with respect to the discriminability of consonant versus vowel features, one-tailed *t* tests revealed that both children and adults actually responded *faster* to /nIk/ than they did to /IEk/ [children = 2,504 msec,  $t(13) = 3.31, p < .005$ ; adults = 1,332 msec,  $t(5) = 2.82, p < .025$ ]. Before interpreting these data, let us consider target responses.

Because adults did not make any false alarms, we discuss only children's target-response percentages. A one-way ANOVA, with stimulus (target vs. foils vs. control) as a within-subjects variable, demonstrated a significant effect of stimulus [ $F(6,78) = 24.93, p < .0001$ ]. Planned, one-tailed *t* tests were again used to address the two main questions: (1) With respect to the featural-overlap hypothesis, children behaved as predicted and made significantly more target responses to /nIk/ (38%) than they did to /gIk/ [21%;  $t(13) = 1.89, p < .05$ ]. They also made marginally more target responses to /nIk/ than they did to /IIf/ [24%;  $t(13) = 1.59, p < .10$ ]; and (2) with respect to the discriminability of single consonant versus vowel features, children made as many target responses to foils that differed from the target by a single vowel feature (/IEk/, 38%) as they did to those differing by a single consonant feature (/nIk/), a pattern different from the

one seen in Experiments 1 and 2. The latter finding is consistent with the reaction-time data, which suggested that /IEk/ was highly confusable with the target for both adults and children.

In summary, the target-response data from Experiments 1–3 demonstrate that the degree of featural overlap, at least for consonant features, is a consistent determinant of children's target-foil confusions (also see Graham & House, 1971). Because featural overlap has proven to be important in studies of adult spoken-word recognition (Goldinger et al., 1989; Luce et al., in press; Milberg et al., 1988), our findings suggest that adults and children may engage in similar lexical recognition processes. However, the data do not reveal as consistent a picture of the relative discriminability of consonant and vowel features. In the first two experiments, foils differing by a single consonant feature were more confusable with the target than were those differing by a single vowel feature. In contrast, the reaction-time and target-response data from Experiment 3 suggest that items differing by a single vowel feature are at least as confusable as those differing by a single consonant feature. Because there were differences in the stimuli across the experiments, including in number of syllables and whether the target formed minimal pairs with other familiar words, these factors should be examined in future research on the discriminability of consonant versus vowel features.

Although the featural-overlap hypothesis appears to provide a good account of the data from Experiments 1–3, two important questions remain, one of which concerns whether the degree of featural overlap between a foil and a target is a determinant of children's target-foil confusions, regardless of the syllable position in which the distinguishing phonetic segments appears. Across all three experiments, foils differing from the target by a single feature on a word-initial segment received more target responses than foils differing by more than one feature. That is, /nItI/ and /nIk/, which each differed from the target by a single feature on the first segment, consistently received more target responses than did foils differing from the targets by two features. Given the inconsistency across the experiments concerning the importance of vowel features, we cannot say whether the featural-overlap account extends to the vowel position in a syllable. Therefore, to test the generalizability of the featural-overlap hypothesis, it would be desirable to determine whether the degree of featural overlap between foil and target in *coda* position also affects children's target-foil confusions. For example, is /Ilg/, which differs from *lick* by one feature on the coda, more confusable with the target than /IIf/, which differs by two features on the coda? This question is important because studies testing children on other tasks have shown that children are better able to detect changes in word-initial segments than in word-final segments (Cole, 1981; Stanovich et al., 1984; Walley, 1987).<sup>7</sup> We addressed this issue in Experiment 4.

The second question left unresolved by Experiments 1–3 concerns the implication of the featural-overlap ac-



count for all versions of the holistic hypothesis. The fact that children appear to be influenced by the featural similarity of target and foils rules out those holistic accounts in which children represent words in terms of overall acoustic or prosodic shape. However, recall that one version of the holistic hypothesis posits that children employ lexical representations composed of phonetic features that are not bundled into segments (e.g., Charles-Luce & Luce, 1990; Logan, 1992); in contrast, adults appear to have access to both featural and segmental representations, as evidenced by their ability to read and spell (see Logan, 1992). The data from Experiments 1–3 do not allow us to determine whether, in addition to featural overlap, segmental overlap also has an effect on children's performance. One way of determining experimentally whether children's lexical representations include features bundled into segments is to test whether changes to more than a single segment are more perceptible than the same aggregate of featural change to a single segment. We also addressed this issue in Experiment 4.

#### EXPERIMENT 4

Experiment 4 had two purposes. First, it sought to determine whether single-feature differences have an equal perceptual effect, regardless of syllable position. To this end, a new foil (/Ilg/) was added, that differed from the target *lick* by a single feature on the coda. According to the featural-overlap hypothesis, both foils that differ from the target by a single consonant feature (i.e., /nlk/ and /Ilg/) should receive more target responses than the foils differing by more than one feature.

Second, Experiment 4 tested the features-and-segments hypothesis, according to which, children's performance in our task is based not only on the number of features that are different between a foil and target, but also on the number of *segments* that are different. In particular, we asked whether a foil that differs from the target by two features on one segment would receive more target responses (false alarms) than a foil that differs by one feature on each of two segments. For example, given the target *lick*, will /gIk/, which differs by two features on the onset, be heard as more similar to the target than /zIlg/, which differs by one feature on the onset and one feature on the coda? If so, we have reason to claim that, contrary to all versions of the holistic hypothesis, both features *and* segments play a role in children's spoken-word representations. To test the features-and-segments hypothesis, two additional foils were created that differed from the target word *lick* by a single feature on each of two segments (see Table 1). The features-and-segments hypothesis predicts that children will make more target responses to the two foils differing from the target by two features on one segment (i.e., /gIk/ and /IIf/) than they will to the two foils differing from the target by one feature on each of two segments (i.e., /zIlg/ and /zEk/). In contrast, the version of the holistic hypothesis in which children represent features that are not

bundled into segments predicts that these four stimuli will receive an equal number of target responses.

#### Method

**Subjects.** The subjects were 15 three- to four-year-old children (8 males and 7 females, age range = 45–53 months, mean = 51). Four additional children were tested, but failed to meet the pretest criterion.

**Materials.** The target word was *lick*, as in Experiment 3. The seven foils and their featural relation to the target are shown in Table 1. The same four control stimuli that were used in Experiment 3 were used.

#### Results and Discussion

Because the number of target-responses was the most informative measure in Experiments 1–3, we focused on it in Experiment 4 (see Table 4). A one-way ANOVA, with stimulus (target vs. foil vs. control) as a within-subjects variable, demonstrated a significant effect of stimulus [ $F(8,112) = 14.10, p < .0001$ ]. This analysis was followed up with two sets of planned comparisons, one testing the predictions of the featural-overlap hypothesis, and the other testing the predictions of the features-and-segments hypothesis.

The featural-overlap hypothesis predicts that children should make more target responses to foils differing from the target by a single feature than to foils differing by two features. Planned *t* tests compared responses to the three foils that differed from the target by a single feature with responses to the four foils that differed by two features. The difference was not significant [ $t(14) = 1.21, p < .15$ ]. A planned *t* test also compared /Ilg/ and /IIf/, to test the more specific prediction of the featural-overlap hypothesis, that foils differing by a single feature in coda position would receive more target responses than foils differing by two features in coda position. Again, the difference was not significant [ $t(14) = 1.10, p < .15$ ]. Unfortunately, the lack of a significant effect of one as opposed to two coda features, taken together with a lack of effect for onset features, makes it impossible to compare the results from Experiment 4 with those from previous studies in which children were better able to detect changes to initial segments than to final segments (Cole, 1981; Stanovich et al., 1984).

Although the predictions of the featural-overlap hypothesis were not supported in the statistical analyses, the ordering of children's target responses generally followed the predicted pattern. Thus, two of the three foils that differed from the target by a single feature (/IEk/ and /Ilg/) received more target responses than the four foils that differed by two features (/gIk/, /IIf/, /zIlg/, and /zEk/). The one exception to the predicted ordering of target responses was that /nlk/, which differed from the target by a single feature, received fewer target responses than /gIk/, which differed by two features. This is puzzling, given that in Experiment 3, children made more target responses to /nlk/ than they did to /gIk/ for the same acoustic strings. It is also puzzling that the differences in target responses predicted by the featural-

overlap hypothesis did not reach significance, whereas they did in the previous experiments. Experiment 4 used more types and tokens of foils than the other experiments (7 types and 21 tokens vs. 3–5 types and 9–15 tokens, respectively), as well as a larger ratio of controls to foils (21:12 vs. 9:18–12:12). Both of these factors might have biased children to make target responses to the foils, thereby obscuring their sensitivity to particular phonetic contrasts. Perhaps the fact that foils in Experiment 4 differed from the target in all three syllable positions also made the task more difficult.

Turning to a consideration of the data in terms of the features-and-segments hypothesis, this hypothesis predicts that those foils differing from the target by two features on one segment (i.e., /gIk/ and /If/) should be more often confused with the target than should the foils differing from the target by one feature on each of two segments (i.e., /zIg/ and /zEk/). To test this prediction, planned, one-tailed *t* tests were performed comparing the target responses made to /gIk/ and /If/ with those made to /zIg/ and /zEk/. The results showed that /gIk/ was confused with the target significantly more frequently than either /zIg/ [ $t(14) = 2.00, p < .05$ ] or /zEk/ [ $t(14) = 3.60, p < .01$ ]. The foil /If/ was more frequently confused with the target than was /zEk/ [ $t(14) = 2.70, p < .01$ ], but responses to it did not differ significantly from those to /zIg/ [ $t(14) = 1.10, p < .20$ ]. These data are consistent with the view that, as well as features, segments are important in children's spoken-word recognition. They are not consistent with the version of the holistic hypothesis in which children represent features that are not bundled into segments.

## GENERAL DISCUSSION AND CONCLUSION

The data from all four experiments are consistent with the view that degree of featural overlap between two acoustic strings is a major determinant of children's spoken-word discrimination. This view is consistent with adult spoken-word recognition (Goldinger et al., 1989; Luce et al., in press; Milberg et al., 1988), as well as with one previous study using a different methodology with children (Graham & House, 1971). In Experiments 1–3, children made significantly more target responses (false alarms) to foils differing from the target by a single consonant feature than they did to those differing by two features. Although the data from Experiment 4 did not demonstrate this pattern in a statistically reliable way, the data were, for the most part, in the predicted direction. The relatively large number of foil types and tokens and the large ratio of foils to control stimuli in Experiment 4 may have obscured children's sensitivity to some phonetic contrasts. Therefore, future experiments employing this method might restrict the number of foil types and tokens, while still examining theoretically important contrasts.

The most important finding of our research is that children made more target responses to foils that dif-

fered from the target by two features on a single segment than they did to foils that differed by one feature on each of two segments. This pattern of responses is consistent with the view that children not only have a representation of phonetic features, but associate these features with segmental positions in the word. Such a finding is at odds with the holistic hypothesis, and suggests that the representations that children use in spoken-word recognition are more similar to those of adults than the data from other tasks, such as gating, have led us to believe. Unfortunately, the predictions of the features-and-segments hypothesis were not completely supported by the data, in that the difference between /If/ and /zIg/ in Experiment 4 was not a reliable one. Furthermore, the overall pattern of results in Experiment 4 failed to show strong evidence for the featural-overlap hypothesis, which is an integral part of the features-and-segments account. Thus, our conclusions at this point must remain tentative.

Importantly, the task introduced in the present study offers a potentially more fruitful way to examine children's spoken-word representations than have many previous paradigms. Since it does not require children to associate phonetic strings with referents, the task may be less demanding of their memory and information-processing abilities. Furthermore, the fact that children compare stimuli with a remembered target increases the likelihood that their performance is based on a lexical representation of the target in memory, and not only on the physical similarity between stimuli.

Future research using this technique might address a number of other issues, one of which concerns the importance in word recognition of vowel features versus consonant features. In Experiments 1 and 2, which used bisyllabic words, foils differing from the target by a single consonant feature received more target responses than did foils differing by a single vowel feature. In Experiments 3 and 4, which used monosyllabic words, vowel and consonant features appeared to have similar perceptual effects. To determine whether the degree of featural overlap among vowels has an effect similar to the featural overlap among consonants, new studies should use monosyllabic and bisyllabic foils that differ from the target by one or two vowel features. Future research using this task should also ask how much of children's performance is determined by the lexical status of the target and foils. The crucial cases in Experiments 1–4 required children to contrast a nonword foil with a remembered target word. Would children's performance differ in theoretically interesting ways if the foils were also words, or if both target and foils were nonwords?

In summary, our results suggest that, while children's spoken-word recognition becomes more efficient over the course of development, the recognition process and the representations involved may be quite similar to those of adults. Although a great deal of research remains to be done, we believe that the task employed here

provides a useful tool for examining children's spoken-word recognition.

## REFERENCES

- ANISFELD, M. (1984). *Language development from birth to three*. Hillsdale, NJ: Erlbaum.
- 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.), *Handbook of child psychology: Infant development* (pp. 573-687). New York: Wiley.
- BARTON, D. (1976a). Phonemic discrimination and the knowledge of words in children under three years. *Papers and Reports on Child Language Development*, **11**, 61-68.
- BARTON, D. (1976b). *The role of perception in the acquisition of phonology*. Bloomington: Indiana University Linguistics Club.
- BARTON, D. (1980). Phonemic perception in children. In G. H. Yeni-Komshian, J. F. Kavanagh, & G. A. Ferguson (Eds.), *Child phonology* (Vol. 2, pp. 97-116). New York: Academic Press.
- BEST, C. T., McROBERTS, G. W., & SITHOLE, N. M. (1988). Examination of the perceptual reorganization for speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 345-360.
- CHARLES-LUCE, J., & LUCE, P. A. (1990). Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language*, **17**, 205-215.
- CHIAT, S. (1979). The role of the word in phonological development. *Linguistics*, **17**, 591-610.
- 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. *Annals of the New York Academy of Sciences*, **379**, 92-109.
- 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.
- DOLLAGHAN, C. A. (1994). Children's phonological neighbourhoods: Half empty or half full? *Journal of Child Language*, **21**, 257-271.
- EDWARDS, M. L. (1974). Perception and production in child phonology: The testing of four hypotheses. *Journal of Child Language*, **1**, 205-219.
- EILERS, R. E., & OLLER, D. (1976). The role of speech discrimination in developmental sound substitutions. *Journal of Child Language*, **3**, 319-330.
- EIMAS, P. D., MILLER, J. L., & JUSCZYK, P. W. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.), *Categorical perception* (pp. 161-195). New York: Cambridge University Press.
- EIMAS, P. D., SIQUELAND, E., JUSCZYK, P., & VIGORITO, K. (1971). Speech perception in infants. *Science*, **171**, 303-306.
- 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.
- FERGUSON, C. A. (1986). Discovering the units of sound and constructing sound systems: It's child's play. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 36-51). Hillsdale, NJ: Erlbaum.
- FERGUSON, C. A., & FARWELL, C. B. (1975). Words and sounds in early acquisition. *Language*, **51**, 419-439.
- FEY, M. E., & GANDOUR, J. (1982). The pig dialogue: Phonological systems in transition. *Journal of Child Language*, **9**, 517-519.
- GARNICA, O. (1971). The development of perception of phonemic differences in initial consonants by English-speaking children: A pilot study. *Papers & Reports on Child Language Development*, **3**, 1-29.
- GERKEN, L. A. (1994). Sentential processes in early child language: Evidence from the perception and production of function morphemes. In H. C. Nusbaum & J. C. Goodman (Eds.), *The transition from speech sounds to spoken words* (pp. 271-298). Cambridge, MA: MIT Press.
- GERKEN, L. A. (in press). A slip of the tongue approach to language development. In J. Charles-Luce, P. Luce, & J. R. Sawusch (Eds.), *Spoken language: Perception, production and development*. Norwood, NJ: Ablex.
- 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.
- GRAHAM, L. W., & HOUSE, A. S. (1971). Phonological oppositions in children: A perceptual study. *Journal of the Acoustical Society of America*, **49**, 559-566.
- 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.
- JAEGER, J. (1992). Not by the chair of my hinny hin hin: Some general properties of speech errors in young children. *Journal of Child Language*, **19**, 355-366.
- JUSCZYK, P. W. (1982). Auditory versus phonetic coding of speech signals during infancy. In J. Mehler, E. C. T. Walker, & M. Garrett (Eds.), *Perspectives on mental representation* (pp. 361-387). Hillsdale, NJ: Erlbaum.
- JUSCZYK, P. W. (1985). On characterizing the development of speech perception. In J. Mehler & R. Fox (Eds.), *Neonate cognition: Beyond the blooming buzzing confusion* (pp. 199-229). Hillsdale, NJ: Erlbaum.
- 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.
- KAIL, R. (1991). Processing time declines exponentially during childhood and adolescence. *Developmental Psychology*, **27**, 259-266.
- KUHL, P. K. (1987). Perception of speech in early infancy. In P. Salapatek & L. B. Cohen (Eds.), *Handbook of infant perception* (Vol. 2, pp. 275-381). New York: Academic Press.
- LOCKE, J. L. (1988). The sound shape of early lexical representations. In M. D. Smith & J. L. Locke (Eds.), *The emergent lexicon: The child's development of a linguistic vocabulary* (pp. 3-22). New York: Academic Press.
- LOGAN, J. (1992). *A computational analysis of young children's lexicons* (Tech. Rep. No. 8). Bloomington: Indiana University, Speech Research Laboratory, Department of Psychology.
- LUCE, P. A. (1986). *Neighborhoods of words in the mental lexicon* (Tech. Rep. No. 6). Bloomington: Indiana University, Speech Research Laboratory, Department of Psychology.
- LUCE, P. A., GOLDINGER, S. D., & AUER, E. T. (in press). Form-based inhibition priming in spoken word recognition: Further evidence for the neighborhood activation model. *Journal of Memory & Language*.
- LUCE, P. A., PISONI, D. B., & GOLDINGER, S. D. (1990). Similarity neighborhoods of spoken words. In G. Altmann (Ed.), *Cognitive models of speech processing* (pp. 122-147). Cambridge, MA: MIT Press.
- MACKEN, M. A. (1979). Developmental reorganization in phonology: A hierarchy of basic units of acquisition. *Lingua*, **49**, 11-49.
- MARSLÉN-WILSON, W. D. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 576-585.
- MARSLÉN-WILSON, W. D., & WELSH, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.
- MENN, L. (1978). Phonological units in beginning speech. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments* (pp. 157-171). Amsterdam: North-Holland.
- MENN, L. (1980). Phonological theory and child phonology. In G. H. Yeni-Komshian, J. F. Kavanagh, & G. A. Ferguson (Eds.), *Child phonology* (Vol. 1, pp. 23-41). New York: Academic Press.
- MENN, L. (1983). Development of articulatory, phonetic, and phonological capabilities. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 3-50). London: Academic Press.
- MENYUK, P., & MENN, L. (1979). Early strategies for the perception

- and production of words and sounds. In P. Fletcher & M. Garman (Eds.), *Studies in language acquisition* (pp. 49-70). Cambridge: Cambridge University Press.
- MILBERG, W., BLUMSTEIN, S., & DWORETZKY, B. (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, **26**, 305-308.
- MILLER, G. A., & NICELY, P. E. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, **27**, 338-352.
- 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.
- SCHVACHKIN, N. K. (1973). The development of phonemic perception in early childhood (E. Dernbach, Trans.). In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 92-127). New York: Holt, Rinehart & Winston. (Original work published 1941)
- SPENCER, A. (1986). Toward a theory of phonological development. *Lingua*, **68**, 3-38.
- STANOVICH, K. E., CUNNINGHAM, A. E., & CRAMER, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology*, **38**, 175-190.
- STEMBERGER, J. P. (1989). Speech errors in early child language production. *Journal of Memory & Language*, **28**, 164-188.
- STRAIGHT, H. S. (1980). Auditory versus articulatory phonological processes and their development in children. In G. H. Yeni-Komshian, J. F. Kavanagh, & G. A. Ferguson (Eds.), *Child phonology* (Vol. 1, pp. 43-72). New York: Academic Press.
- STUDDERT-KENNEDY, M. (1986). Sources of variability in early speech development. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 58-76). Hillsdale, NJ: Erlbaum.
- STUDDERT-KENNEDY, M. (1987). The phoneme as perceptuo-motor structure. In A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading and writing* (pp. 67-84). Orlando, FL: Academic Press.
- TREIMAN, R., & BARON, J. (1981). Segmental analysis ability: Development and relation to reading ability. In G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 3, pp. 159-198). New York: Academic Press.
- TREIMAN, R., & BREAU, A. M. (1982). Common phoneme and overall similarity relations among spoken syllables: Their use by children and adults. *Journal of Psycholinguistic Research*, **11**, 569-598.
- VIHMAN, M. M. (1993). Variable paths to word production. *Journal of Phonetics*, **21**, 61-82.
- VIHMAN, M. M., & VELLEMAN, S. L. (1989). Phonological reorganization: A case study. *Language & Speech*, **32**, 149-170.
- 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. (1993). The role of vocabulary growth 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., 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.
- WATERSON, N. (1971). Child phonology: A prosodic view. *Journal of Linguistics*, **7**, 179-211.
- WERKER, J. F. (1991). The ontogeny of speech perception. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 91-109). Hillsdale, NJ: Erlbaum.
- WIJNEN, F. (1992). Incidental word and sound errors in children's language production. *Journal of Memory & Language*, **31**, 734-755.

## NOTES

1. Different *types* of consonant features (e.g., place vs. manner of articulation) might also result in different degrees of similarity (e.g., Cole, 1981); however, we do not address that issue in the current research.

2. If the likelihood that children form a segmental representation of a word increases with their knowledge of the word (Walley, 1993; Walley & Metsala, 1990), *little* might well be given a segmental representation. However, the mechanism by which knowledge of a single word (as opposed to knowledge of several phonetically similar words) could lead to a change in lexical representation has not been well specified.

3. We thank an anonymous reviewer for pointing out that the pattern of adult reaction times is consistent with data from mispronunciation-detection tasks (MP tasks), in which listeners respond more slowly to word-initial mispronunciations than they do to word-internal mispronunciations (Cole, 1973). Such data have been taken as evidence for the special status of word-initial segments in lexical access. However, there are four reasons why our data should not be interpreted in this way. First, MP tasks typically conflate word-initial segments with first syllables (e.g., Cole, 1973); therefore, they can be interpreted as indicating that information in the first syllable of a word (but not necessarily its first segment) is crucial for lexical access. Second, the MP task is conceptually very different from our task, in which listeners had to maintain a single word in memory and compare stimulus items to it, and thus certainly could have performed the task without attempting lexical access of the stimuli. In contrast, listeners in the MP task probably attempt lexical access for a stimulus and decide that a word has been mispronounced either when access fails or when a candidate word is found that differs minimally from the stimulus. Third, the reaction-time data from Experiment 3 are not consistent with this view; rather, they are longest for foils differing from the target on word-internal segments. Finally, subjects in the MP task typically make *fewer* false alarms to items with mispronunciations on the initial segment—exactly the opposite pattern to that seen in adults and children in our study. In sum, therefore, it seems unlikely that the adult reaction-time data in Experiment 1 reflect listeners' special attention to word-initial segments.

4. Showing that children's lexical representations are based on features leaves open the possibility that these features are not bundled into segments. We address this point in Experiment 4.

5. Note that while /lɪtn/ shares its first three segments with *little*, there is a glottal stop before the final /n/, thereby making this foil different from the target by at least two features.

6. We thank an anonymous reviewer for pointing out that "nick" is a verb and a name in English. Although we believe that the verb is sufficiently unfamiliar to 3- to 4-year-olds for it to be counted as a non-word for them, it is possible that some of the children had siblings or friends named "Nick." Even if this were the case, it is important to note that children's performance on this item is generally consistent with the overall pattern of data for foils that differed from the target by a single feature.

7. Note that children's special attention to word-initial segments in these studies, but not in others, is consistent with our proposal that differences between children and adults do not necessarily reflect differences in lexical representation, but, rather, they may reflect different approaches to particular experimental tasks.

(Manuscript received September 22, 1993;  
revision accepted for publication November 14, 1994.)