

When Half a Word Is Enough: Infants Can Recognize Spoken Words Using Partial Phonetic Information

Anne Fernald, Daniel Swingley, and John P. Pinto

Adults process speech incrementally, rapidly identifying spoken words on the basis of initial phonetic information sufficient to distinguish them from alternatives. In this study, infants in the second year also made use of word-initial information to understand fluent speech. The time course of comprehension was examined by tracking infants' eye movements as they looked at pictures in response to familiar spoken words, presented both as whole words in intact form and as partial words in which only the first 300 ms of the word was heard. In Experiment 1, 21-month-old infants ($N = 32$) recognized partial words as quickly and reliably as they recognized whole words; in Experiment 2, these findings were replicated with 18-month-old infants ($N = 32$). Combining the data from both experiments, efficiency in spoken word recognition was examined in relation to level of lexical development. Infants with more than 100 words in their productive vocabulary were more accurate in identifying familiar words than were infants with less than 60 words. Grouped by response speed, infants with faster mean reaction times were more accurate in word recognition and also had larger productive vocabularies than infants with slower response latencies. These results show that infants in the second year are capable of incremental speech processing even before entering the vocabulary spurt, and that lexical growth is associated with increased speed and efficiency in understanding spoken language.

INTRODUCTION

Understanding spoken language requires the rapid integration of acoustic information with linguistic and conceptual knowledge. To follow a casual conversation, skilled listeners typically process 10 to 15 phonemes per second and have no trouble recognizing words spoken at twice that rate (Cole & Jakimik, 1980). Many studies with adults have shown that the ability to process speech continuously is central to this remarkable efficiency. By making use of phonetic information incrementally, the listener is able to identify spoken words rapidly, often before their acoustic offset (Marslen-Wilson, 1987; McClelland & Elman, 1986). Although there is wide agreement that the ability to recognize words using partial phonetic information is a robust characteristic of adult speech processing, little is known about the early development of this critical capacity. Current models of adult word recognition either do not address such developmental questions, or make the prediction that incremental speech processing is a skill that develops slowly. For example, Charles-Luce and Luce (1990) and Walley (1993) argued that children make use of word-initial information in lexical access only after they have acquired a substantial lexicon, a view supported by research showing that school-age children are less capable than adults in using partial phonetic information to identify familiar words. The goal of the present research was to evaluate these claims, investigating whether very young children are able to identify familiar words in fluent speech based on partial pho-

netic information, and whether the early development of this important speech processing capacity is related to lexical development.

According to current psycholinguistic theories of spoken word recognition, listeners generate and reject hypotheses about word identity based on what they have heard up to that moment (Marslen-Wilson & Zwitserlood, 1989). For example, a word onset such as /*per*/ activates several English words including *peruse*, *periphery*, *perimeter*, and others consistent with the initial phonetic information. When the listener hears /*perl*/, most of these candidates can be eliminated; then, when the /*m*/ is heard, the word *perimeter* can be uniquely identified before the final syllable is spoken. Although words initially rejected can be reactivated if later information matches well (e.g., *perimeter* may be activated by *terimeter*; Allopenna, Magnuson, & Tanenhaus, 1998; Connine, Blasko, & Titone, 1993), it is clear that attention to word-initial information facilitates rapid decisions about spoken word identity. Another advantage of continuous speech processing is that incremental lexical activation may help the listener to segment utterances into words. In computational models of word recognition, candidate words compete to account for stretches of speech as a sentence unfolds. The parse finally selected, if correct, will be the parse that assigns each bit of speech to a word with no speech unac-

counted for (Norris, McQueen, Cutler, & Butterfield, 1997). In this way, lexical activation itself provides one solution to the segmentation problem, supplemented by other language-specific cues to word boundaries. If such mechanisms enhance the efficiency of speech processing for adults, they could be especially beneficial to infants first learning a language.

Although the ability to recognize words using word-initial information is central to competence in spoken language understanding (Grosjean, 1985), several studies suggest that children only slowly develop this skill. When Cole (1981) and Walley (1987) asked adults and kindergartners to detect mispronunciations occurring in either word-initial or word-final position, adults performed much better on word-initial mispronunciations. Kindergartners were less accurate overall, and only showed this word-initial bias when the words were contextually highly constrained. Consistent with these findings, Walley (1988) used a gating technique to show that although children could recognize word-initial fragments, they needed relatively more phonetic information than adults to identify partial words correctly. These results were interpreted as evidence that children do not process the words they hear continuously, perhaps because they initially have more "holistic" lexical representations that lack segmental structure. In this case, it might be advantageous for children to wait until all available phonetic information has been heard before identifying a word. Charles-Luce and Luce (1990) have also argued that because young children's lexicons are steadily expanding, premature commitment based on partial information could lead to frequent errors and interfere with word learning.

Two recent studies of spoken language understanding by infants in the second year of life, however, suggest that these earlier findings may have underestimated the ability of young children to make use of the phonetic information at the beginnings of familiar words. Using new experimental methods that provide detailed information about the time course of early word recognition, Fernald, Pinto, Swingley, Weinberg, and McRoberts (1998) assessed the speed and accuracy of word recognition by infants at 15, 18, and 24 months of age. Infants' eye movements were measured as they looked at pictures of familiar objects while listening to speech labeling one of the pictures. Fernald et al. found that between the ages of 15 and 24 months, infants become much faster at and more reliable in identifying familiar spoken words. Although the 15-month-olds tended to shift their gaze from the distracter picture to the labeled target picture only after hearing the entire tar-

get word, the 24-month-olds were more likely to initiate a shift before the target word was completely spoken.

Although the findings of Fernald et al. (1998) suggest that 2-year-olds do not need to wait until after the acoustic offset of the word before correct identification can occur, it was not clear from these initial results just how much phonetic information infants required before recognizing the target word. Swingley, Pinto, and Fernald (1999) addressed this question more directly by presenting 24-month-olds with pictures of objects whose names either overlapped phonetically at word onset (e.g., *doggie* and *doll*) or did not overlap (e.g., *doggie* and *tree*). Whereas *doggie* was potentially distinguishable from *tree* at the beginning of the word, *doggie* and *doll* did not diverge until the occurrence of the second consonant, about 300 ms into the signal. Swingley et al. found that 24-month-olds responded rapidly when distinguishing *doggie* from *tree* but delayed their response by about 300 ms when distinguishing *doggie* from *doll*. This finding indicates that young children are indeed capable of monitoring speech continuously and attending to word-initial information in identifying words. The discrepancy between these positive findings and the more negative picture emerging from earlier studies is most likely the result of substantial differences in task demands. The procedures used by Walley (1987, 1988) and others required children to make explicit judgments about the identity of spoken words in tasks that could have been somewhat confusing to them, and these may have underestimated children's speech processing capabilities. In contrast, the visual preference procedure used in our research takes advantage of the tendency of infants to look spontaneously at pictures when named with a familiar word, an engaging and less demanding task that is also more representative of children's daily experience (see Swingley, Pinto, & Fernald, 1998).

The Swingley et al. (1999) study showed that 2-year-old children, like adults, can make use of phonetic information as they hear it when identifying spoken words. It is not possible to conclude from these findings, however, that hearing the first portion of the word was sufficient for word recognition as it often is for adults, because the infants' responses did not occur at exactly the point when the target word diverged from the name of the distracter. For example, the duration of the target word *doggie* was 973 ms, and the mean response latency to *doggie* when paired with *doll* was 1,017 ms. Although the shift in gaze was sometimes initiated before the end of the word, on average it occurred about 600 ms after the beginning of the /g/, the point at which the two words

became phonetically distinct. Of course, the response latency includes not only the time required to identify the target word, but also the time required to initiate a shift in gaze, and only rough estimates can be made about the amount of time it takes for oculomotor programming to occur in this situation. This lag in response time, however, is consistent with the possibility that infants continued to listen for a few hundred milliseconds beyond the point at which the word became uniquely identifiable, rather than responding as soon as disambiguating information became available. If infants can identify spoken words when only the word-initial information is presented, this would provide more conclusive evidence for the early development of incremental processing. The present study extended previous findings by asking whether infants younger than 24 months are able to identify words when only the first part of the word is presented.

A second question motivating this research was whether speed and accuracy in spoken word recognition are related to growth in productive vocabulary. In particular, is the use of word-initial phonetic information in word recognition possible only for children who are relatively more advanced in vocabulary development, as Walley (1987, 1988) and Charles-Luce and Luce (1990) have suggested? Toward the end of the second year, infants typically enter a period known as the "vocabulary burst," during which the rate of learning new words begins to accelerate. After gradually acquiring their first 50 words or so over a period of several months, they increase the pace and begin to learn several new words a week (Bloom, 1993; Fenson et al., 1994). Numerous studies have explored the relation of the vocabulary burst to other developmental changes in linguistic and cognitive competence around this age that might account for an increase in the facility of word learning. One approach has been to investigate how growth in productive vocabulary size is related to growth in receptive vocabulary (Bates, Bretherton, & Snyder, 1988; Harris, Yeles, Chasin, & Oakley, 1995). Other studies have focused on the development of conceptual abilities that are not specific to language processing which could be prerequisite for progress in linguistic development (Gopnik & Meltzoff, 1987). Bates and MacWhinney (1987) and Bloom (1993) have also emphasized the importance of understanding how developmental changes in a wide range of information-processing skills may influence the early stages of acquiring a lexicon. There is, however, little evidence that directly links growth in vocabulary to the emergence of speech processing skills that are clearly related to competence in understanding spoken lan-

guage. The present study explored the connection between vocabulary growth and speech comprehension by relating parental-report measures of productive vocabulary size to "on-line" experimental measures, which monitor the time course of word recognition as it occurs.

To determine whether children younger than 2 years of age are able to process speech continuously, 21-month-olds (Experiment 1) and 18-month-olds (Experiment 2) were tested in the on-line word recognition procedure developed in earlier research (Fernald et al., 1998; Swingley et al., 1999). Infants were presented with familiar object names in their intact form and also as word fragments consisting of only the first two or three phonemes of the word. At each age, infants' reliability and speed of response in identifying whole and partial target words in continuous speech were assessed. After combining the data from Experiments 1 and 2, analyses were then performed to assess whether those infants who were more advanced in vocabulary development were also more efficient in spoken word recognition.

EXPERIMENT 1

Method

Participants. Participants were 16 male and 16 female ($N = 32$) 21-month-old infants ($range = 624-653$ days). An additional 6 infants participated in the study but were not included in the final sample for the following reasons: failure to complete the test session because of fussiness or inattentiveness (2), parental interference during testing (1), no vocabulary checklist submitted (2), and experimenter error (1). All participants were from families in which English was the primary language spoken.

Auditory stimuli. The speech stimuli were sentences containing one of four target words (*doggie, baby, birdie, kitty*); all four words were understood by the participating infants according to parental report. Each target word occurred in the final position in the same two carrier phrases: "Where's the [target]?" and "Can you see the [target]?" An English-speaking woman recorded several tokens of each pair of sentences with comparable intonation contours but slightly varied speaking rates. These candidate stimulus sentences were recorded using a Revox B77 tape recorder, and then digitized, analyzed, and edited using the SoundEdit 16 (1996) waveform editor on a Macintosh computer. Acoustic measurements of the duration of the carrier phrase and target word in each token were used to select the best matched tokens for the final stimulus set. The mean duration of the "Where's the . . ." carrier phrase across target words

was 605 ms (*range* = 574–674 ms); the mean duration of the target word was 725 ms (*range* = 662–883 ms). Tokens were matched on the basis of the first carrier phrase only because it was the infant's response to the first presentation of the target word in each trial that was of interest. Each target word sentence was used in both the whole-word and partial-word conditions. Partial-word sentences were constructed by eliminating the final portion of the target word in each whole-word sentence, leaving only the initial consonant and first part of the vowel. The partial-word stimuli were: /*bei*/ (*baby*), /*daw*/ (*doggie*), /*ki*/ (*kitty*), and /*ber*/ (*birdie*). Through careful selection of tokens and editing of the waveforms, it was possible to make all the partial target words comparable in duration ($M = 309$ ms, *range* = 291–339 ms).

In addition to the target word sentences, "ostensive" and filler sentences were also recorded. For the ostensive sentences, each target word was recorded in the same carrier phrase ("That's a [target]") for use in the familiarization trials preceding the test trials. Target words in the ostensive sentences were presented only as whole words and were not matched for duration. Four filler sentences were also recorded with the object names *monkey*, *teddy*, *apple*, and *shoe*. Since filler trials were included only to maintain infants' interest by providing visual and auditory variation, these were presented as whole words only and each had a different carrier phrase.

Visual stimuli. Visual stimuli consisted of brightly colored, digitized images of objects corresponding to the target words, taken from photographs in children's picture books. For each target word, two different object tokens were used. Images were approximately matched in size and brightness, and were presented on 25 cm × 19 cm computer monitors.

Apparatus. The experiment was conducted in a sound-treated room containing a booth constructed of three panels covered in cloth and open on the fourth side. The side panels of the testing booth measured 1 m × 2 m, and the front panel measured 1 m × 1.2 m. The two computer monitors and a loudspeaker were mounted in the front panel. During testing the infant sat on the parent's lap facing the monitors. A curtain suspended between the two sides of the booth hung down behind the infant's head, obstructing the parent's view of the monitors but allowing the infant access to the parent during the test session. The monitors were positioned at the infant's eye level 60 cm apart horizontally; the loudspeaker was located on the floor centered between the monitors. A video camera mounted behind the front panel focused on the infant's face. The camera was connected to a video recorder in an adjacent control

room, where the computer controlling the experiment was also located.

Procedure. During a 15-min familiarization period prior to the test session, an experimenter explained the procedure to the parent and interacted with the child. At this time the parent filled out the consent form and submitted the MacArthur Communicative Development Inventory (CDI: words and sentences subscale) form, which had been completed at home. When parent and child appeared to be comfortable, they were seated in the testing booth. The lights in the room were dimmed as two identical pictures of toy ducks appeared on the two computer monitors in the booth. A second experimenter located in the control room spoke briefly to the child over the loudspeaker, to acquaint the child with the sound source and the positions of the monitors. When the child was relaxed and attentive, the experimental session began.

The first four stimulus presentations consisted of familiarization trials. On these initial trials, a single picture was presented either on the left or the right accompanied by an ostensive sentence naming the object. Following the familiarization trials, 20 test trials were presented: 8 whole-word trials, 8 partial-word trials, and 4 filler trials. Trial types were presented in a quasi-random order, with each object appearing as target and as distracter an equal number of times in whole- and partial-word trials for each participant. Side of presentation of target and distracter objects was counterbalanced within each condition. Four stimulus orders were created by reversing the order of test trials front to back and then reversing the sides of target presentations within each of these orders. Each order was presented to an equal number of boys and girls. On each test trial the two pictures were shown in silence for 3 sec prior to presentation of the speech stimulus. The test phase of the trial was the period beginning at the onset of the target word in the first sentence. During the 1-sec intertrial interval the monitor screens were black. The overall duration of the 24-trial sequence was about 4 min.

Coding. During the testing session a digital time code accurate to 33 ms was recorded onto the videotape, as well as a visual marker indicating the onset of the speech stimulus on each trial. Trained observers, blind to the side of the target picture, coded the time course of the session off-line in a frame-by-frame analysis of the videotape, noting on each frame whether the infant's eyes were oriented toward the picture on the left, toward the picture on the right, between the pictures, or away from both pictures. These data on the time course of the infant's eye movements in response to the speech signal were then entered into a computer and aligned with the onset of the target word on each trial. The computer

then calculated the duration of each look and indicated the time at which the infant initiated each shift in gaze. Inter- and intraobserver reliability checks were conducted routinely for all coders. For 25% of the participants, randomly selected blocks of three trials were coded independently by two different coders; the response latencies measured by these coders differed by zero or one frame (33 ms) for 96.8% of the latencies.

Accuracy measures. To determine the accuracy of infants' responses to whole and partial target words, it was first necessary to specify what should count as a correct response. Defining criteria for accuracy in an eye movement experiment with infants is less straightforward than with adults. Whereas adults can be instructed to fixate on a central point at trial onset, to look at the target object only when they hear the relevant speech stimulus, and to remain on the target until the end of the trial, infants cannot be constrained in this way. Thus the behavior that constitutes a correct response varies with the position of the infant's eyes at the onset of the target word. If the infant by chance is looking at the distracter picture when the target word is spoken, the correct response is to shift fixation immediately to the target picture; if the infant is already looking at the target picture at target word onset, the correct response is to continue looking at that picture and not to shift immediately. Trials were pooled across participants and sorted according to where the infant was looking at the onset of the target

word in the first stimulus sentence in each trial. Consistent with previous research (Fernald et al., 1998; Swingley et al., 1999) the following distribution was found: on 45% of trials the infant initially fixated on the distracter (distracter-initial trials); on 42% of the trials the infant initially fixated on the target (target-initial trials); and on 13% of the trials the infant looked away from both monitors (away trials).

Figure 1A shows the distribution of response times on distracter-initial trials in the whole-word condition. Infants shifted to the target before the end of the trial on 82% of the distracter-initial trials; on 12% of these trials they never shifted to the target, remaining on the distracter throughout the trial. Rather than accept any shift to the target picture following target word onset as a correct response, some very fast and very slow responses were excluded from consideration.

First, trials on which the shift occurred within the first 366 ms following target word onset were eliminated, on the assumption that these shifts had been initiated before the word was heard. This cutoff was based on previous research on infant eye movements. Canfield, Smith, Brezsnayak, and Snow (1997) and Haith, Wentworth, and Canfield (1993) estimated that the minimum latency for 3-month-old infants to initiate a shift in fixation to a peripheral stimulus is from 133 to 200 ms. Using a different visual attention task with older infants, Hood and Atkinson (1993) found that responses were further delayed by 200 ms when

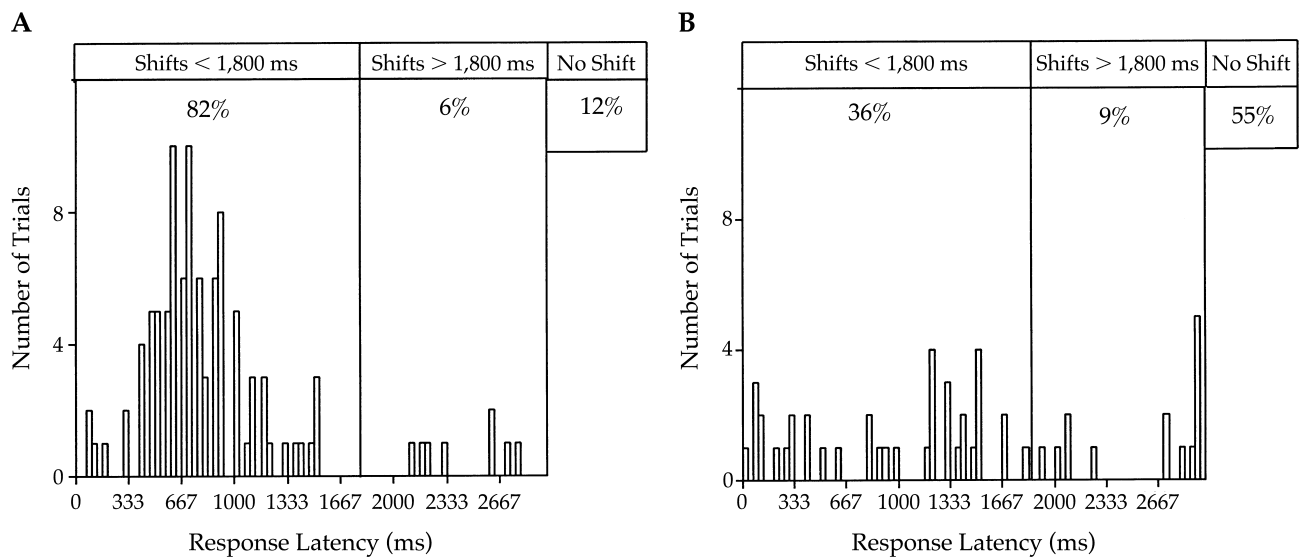


Figure 1 Distribution of response latencies (in ms) for the first shift from the initially fixated picture to the alternative picture on 443 whole-word trials, following onset of the spoken target word. (A) Distracter-initial trials, on which the correct response was to shift from the distracter picture to the target picture within 1,800 ms of target word onset. (B) Target-initial trials, on which the correct response was to maintain fixation on the target picture for 1,800 ms or longer. The proportion of trials on which no shift occurred within 3,000 ms is noted on the right.

infants had to disengage from one stimulus before initiating a shift in fixation to a second stimulus. Because the procedure used in the present study required disengagement from one picture before shifting to the other, the lower cutoff of 366 ms seemed reasonable.

Second, only shifts occurring within 1,800 ms of the first target word onset were counted as correct. This criterion was chosen to eliminate responses influenced by the second repetition of the target word, and excluded only outliers more than 2 *SDs* greater than the mean of the distribution.

Figure 1B shows the distribution of response latencies on the 200 target-initial trials in the whole-word condition. The criterion for correct responses on these trials was complementary to the criterion for distracter-initial trials: the response was considered correct only if the infant continued to fixate on the target for at least 1,800 ms.

Results

The first question of interest was whether 21-month-old infants were able to recognize partial as well as whole target words. Overall accuracy scores were calculated for each infant based on the percentage of correct responses for each trial type. Correct trials included distracter-initial trials on which the infant shifted to the target picture between 366 and 1,800 ms from target word onset, and target-initial trials on which the infant stayed on the target picture for 1,800 ms or longer. If infants responded randomly in this procedure, their looking behavior on each trial would be unrelated to the spoken target word; that is, infants would be just as likely to continue looking at the distracter picture as at the target picture, and just as likely to shift to the other picture on target-initial trials as on distracter-initial trials. In this case the overall accuracy score would be around 50%. The first hypothesis—that accuracy would be greater than expected by chance on both whole- and partial-word trials—was confirmed. Mean accuracy scores were significantly above 50% for both whole words ($M = 76.53\%$, $SD = 14.96\%$), $t(31) = 7.97$, $p < .001$, and partial words ($M = 70.05\%$, $SD = 23.78\%$), $t(31) = 4.77$, $p < .001$.

The finding that 21-month-old infants were successful in recognizing partial words and whole words led to the next question: were whole words recognized more quickly and/or more reliably than partial words? When listening to whole-word stimuli, 21-month-olds responded with a mean latency of 749.81 ms ($SD = 192.82$ ms) on distracter-initial trials; when listening to partial-word stimuli, the mean response

latency was 828.44 ms ($SD = 292.35$ ms). Mean percent correct scores and mean response latencies were analyzed in two-way analyses of variance (ANOVAs), with word type (whole versus partial) as the repeated measure and gender as the grouping variable. The analysis of accuracy scores revealed no significant main effects or interaction. The analysis of reaction times also showed no significant differences. These results show that 21-month-olds responded as rapidly and reliably to partial-word stimuli as they did to whole-word stimuli.

EXPERIMENT 2

The second experiment investigated whether 18-month-old infants would also be successful in recognizing familiar words using only partial phonetic information. Previous research (Fernald et al., 1998) found that 18-month-olds who heard intact familiar words typically waited until after the end of the target word before shifting their gaze to the named picture. This result could indicate that infants at this age need to hear all or most of the word before lexical access can reliably occur, and thus that 18-month-olds would not perform as well as 21-month-olds on partial-word trials.

Method

Participants. Participants were 16 female and 16 male ($N = 32$) 18-month-old infants ($range = 533$ – 564 days). An additional 7 infants participated in the study but were not included in the final sample for the following reasons: failure of the infant to complete the test session because of fussiness or inattentiveness (4); no vocabulary checklist submitted (2); and experimenter error (1). All participants were from families in which English was the primary language spoken.

Auditory stimuli. Because many parents of the 18-month-olds in this experiment's population reported that their children did not yet know the words *birdie* and *kitty*, it was necessary to replace these two target words. Therefore, the target words used in Experiment 2 were *baby*, *doggie*, *car*, and *ball*—all familiar to the 18-month-olds according to parental report. The partial-word stimuli were /*bei*/ (*baby*), /*daw*/ (*doggie*), /*ka*/ (*car*), and /*baw*/ (*ball*).

Visual stimuli. The visual stimuli consisted of the same pictures of babies and dogs used in Experiment 1; pictures of birds and cats were replaced with pictures of cars and balls to match the new target words.

Procedure. The apparatus, procedure, and coding methods in Experiment 2 were the same as described for Experiment 1.

Results

In Experiment 2 it was found that 18-month-olds were able to identify word fragments as well as intact words. Accuracy was significantly above 50% for both whole words ($M = 75.39\%$, $SD = 14.96\%$, $t(31) = 9.60$, $p < .001$, and partial words ($M = 68.98\%$, $SD = 19.00\%$, $t(31) = 5.65$, $p < .001$). The mean response latency was 943.31 ms ($SD = 204.03$ ms) on whole-word trials and 862.48 ms ($SD = 227.89$ ms) on partial-word trials. Percentage of correct scores and response latencies were analyzed in two-way ANOVAs, with word type (whole versus partial) as the repeated measure and gender as the grouping variable. The analysis of accuracy scores revealed no significant main effects or interaction. The analysis of reaction times also showed no significant differences. As with the 21-month-old infants in Experiment 1, the 18-month-olds in Experiment 2 responded as rapidly and reliably to partial-word stimuli as to whole-word stimuli.

SPEED AND ACCURACY OF WORD RECOGNITION IN RELATION TO AGE AND VOCABULARY SIZE

The next series of analyses was run to determine whether efficiency in spoken word recognition was related to level of vocabulary development and age. As is typical for children toward the end of the second year, there was considerable variability in productive vocabulary size with a third of the younger infants having higher CDI scores than a third of the older infants. There was a gap in the distribution of CDI scores, with 26 infants speaking fewer than 60 words and 38 infants speaking more than 100 words. Many investigators have observed that at the beginning of the vocabulary spurt, infants typically have around 40 to 70 words in their productive vocabulary (Bloom, 1993). For this reason, the samples at both ages were divided into two groups, one with lower vocabulary scores (<60 words) and the other with higher vocabulary scores (>100 words). Although a measure for rate of word learning was not available, it was possible to infer from the norms that infants in the high-vocabulary group were more likely to have begun the vocabulary spurt than infants in the low-vocabulary group. To investigate whether accuracy and speed varied with vocabulary size at each age level, the percentage of correct scores and response latencies for 21-month-olds (Experiment 1) and 18-month-olds (Experiment 2) were each analyzed in two-way ANOVAs. Word type (whole versus partial) was the within-subjects variable, and vocabulary size (<60 words versus >100 words) was the grouping variable.

Accuracy in word recognition was found to vary with productive vocabulary size for both 21- and 18-month-old infants. At 21 months, infants in the high-vocabulary group ($M = 75.59\%$, $SD = 20.62\%$) performed significantly better than infants in the low-vocabulary group ($M = 57.16\%$, $SD = 22.09\%$), $F(1, 30) = 5.05$, $p < .05$. At 18 months as well, the high-vocabulary group ($M = 77.79\%$, $SD = 15.09\%$) performed significantly better than the low-vocabulary group ($M = 69.64\%$, $SD = 17.75\%$), $F(1, 30) = 4.32$, $p < .05$. There were no other significant main effects or interactions at either age. Thus, at both ages, those infants who spoke more words were also more accurate in understanding than were those who had smaller speaking vocabularies.

The analyses of mean reaction times within each age group were less consistent than the accuracy measures, as shown in Figure 2. At 21 months, the high-vocabulary infants ($M = 780.21$ ms, $SD = 240.36$ ms) had faster reaction times than the low-vocabulary infants ($M = 920.67$ ms, $SD = 240.92$ ms), although the difference was not reliable, $F(1, 24) = 1.26$, $p > .10$. At 18 months, the pattern was reversed: the low-vocabulary group ($M = 850.19$ ms, $SD = 190.65$ ms) had significantly faster reaction times than the high-vocabulary group ($M = 1010.23$ ms, $SD = 240.76$ ms), $F(1, 23) = 5.40$, $p < .05$. The interaction of vocabulary size with word type was also significant, $F(1, 23) = 4.70$, $p < .05$, indicating that the response speed difference between the high- and low-vocabulary groups was limited to the partial-word trials. When the comparison focused on whole-word trials only, however, mean reaction times for high-vocabulary infants were faster than for low-vocabulary infants at both 18 months ($M = 932.66$ ms versus 957.43 ms) and 21 months ($M = 801.97$ ms versus 826.55 ms), although the main effect for vocabulary size was not significant.

After examining each age group separately, the data from Experiments 1 and 2 were combined into one sample to determine whether accuracy and speed varied with age as well as with vocabulary size. The mean percentage of correct scores were analyzed in a three-way ANOVA with word type (whole versus partial) as the repeated measure, and age (18 versus 21 months) and vocabulary size (<60 words versus >100 words) as the grouping variables. In the combined sample, the mean accuracy score for high-vocabulary infants was 76.17% ($SD = 19.25\%$), whereas the mean for low-vocabulary infants was 67.72% ($SD = 18.79\%$). The main effect for vocabulary size was significant, $F(1, 60) = 9.61$, $p < .01$, consistent with the previous analyses of accuracy scores, with no effects for age or word type and no interactions. In a comparable analysis of response times, there were no signifi-

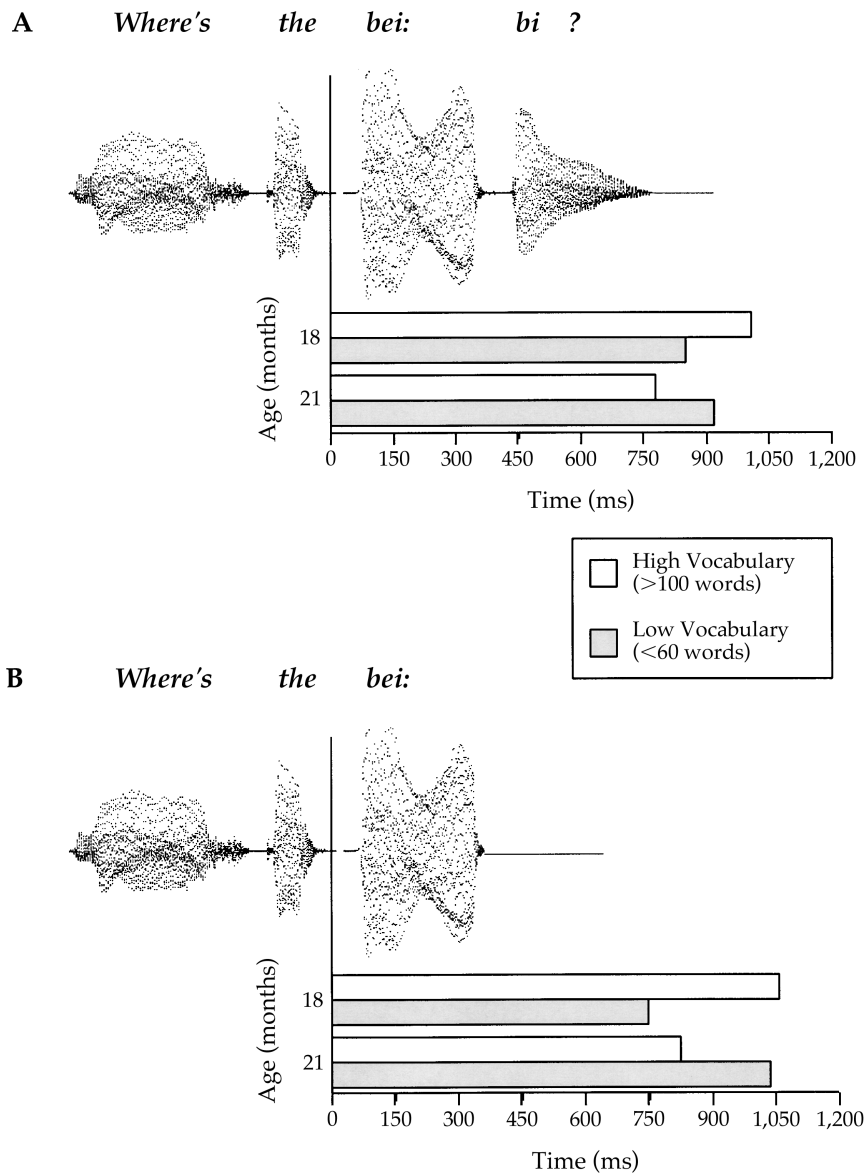


Figure 2 Mean response latencies (in ms) on (A) whole-word trials and (B) partial-word trials for infants in the low-vocabulary (<60 words) and high-vocabulary (>100 words) groups at 18 and 21 months of age. This analysis included only those trials on which infants initially looked at the distracter picture and then quickly shifted to the target picture when the target word was spoken. The graphs are aligned with waveforms of one of the stimulus sentences in both whole- and partial-word forms.

inant main effects for age or vocabulary size, although the Age \times Vocabulary Size interaction was reliable, $F(1, 47) = 4.52, p < .05$. This interaction reflected the fact that among the 18-month-olds, the mean overall reaction time was higher for high-vocabulary infants than for low-vocabulary infants, whereas the reverse pattern prevailed for the 21-month-olds (see Figure 2).

Although larger vocabulary size was expected to be associated with faster speed as well as with greater accuracy, the findings on the relation of vocabulary to reaction time were inconsistent. One likely reason for

this inconsistency is that 18% of the participants could not be included in the analysis of response latencies. Although mean accuracy scores were calculated for all infants based on performance across all trials, mean reaction times could only be calculated for infants who shifted correctly to the target picture on at least one distracter-initial trial. Reflecting the fact that infants were somewhat less accurate on partial- than on whole-word trials, 12 participants failed to shift appropriately on any distracter-initial trial in the partial-word condition, compared with 1 participant in the

whole-word condition. Because reaction times were not available in both conditions for these 13 infants, they could not be included in repeated-measures ANOVAs comparing response speeds on whole- and partial-word trials. Moreover, the missing participants comprised a nonrandom subset of the total sample, consisting of infants with poorer performance overall. Of the 18 participants whose mean reaction times on whole-word trials were slower than 1,000 ms, 9 failed to shift correctly on any partial-word trial. In contrast, of the 45 infants with mean reaction times faster than 1,000 ms, only 3 failed to shift correctly on any partial-word trial. The 13 infants omitted from the reaction time analysis were also less accurate as well as slower than those included, scoring only 61.5% correct on whole-word trials, on average, compared with a mean of 79.6% correct for the 51 infants with complete reaction time data.

Because reaction time data were missing for so many participants in the partial-word condition, the relation of vocabulary size to speed and accuracy of word recognition was explored further by focusing on response latencies in the whole-word condition. Noting that 22 of the 38 infants in the high-vocabulary group had mean reaction times above the median, whereas only 9 of 25 infants in the low-vocabulary group were above the median, infants were classified by average speed of response as well as by vocabulary size. The sample was divided at the median into two groups, a fast group with mean reaction times of 517 to 860 ms, and a slow group with mean reaction times of 867 to 1,389 ms. Infants in the fast group were found to have significantly more words in their productive vocabulary ($M = 245.68$, $SD = 171.25$) than infants in the slow group ($M = 142.62$, $SD = 136.63$), $F(1, 61) = 6.99$, $p < .05$, as shown in Figure 3.

Grouping the infants by mean reaction times as well as CDI scores also made it possible to examine response accuracy in relation to response speed and vocabulary size. The accuracy scores on whole- and partial-word trials were examined in a three-way ANOVA with word type as the repeated measure, and response speed (fast versus slow) and vocabulary size (>100 words versus <60 words) as the grouping variables. Infants in the fast group ($M = 77.83\%$, $SD = 18.48\%$) were significantly more accurate than infants in the slow group ($M = 68.51\%$, $SD = 18.52\%$), $F(1, 59) = 5.86$, $p < .05$. As noted previously, infants with larger productive vocabularies performed more accurately than those who spoke fewer words, although the main effect for vocabulary size was only marginally significant in this analysis, $F(1, 59) = 3.12$, $p < .09$. This analysis also revealed a significant main effect for word

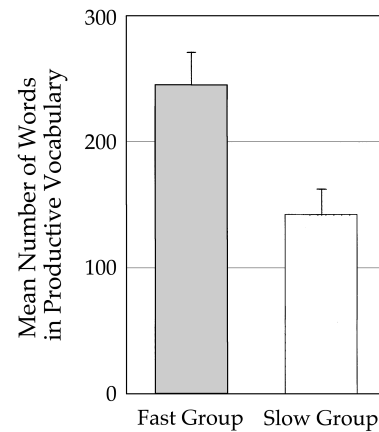


Figure 3 Mean numbers of spoken words in the vocabularies of infants in the fast and slow groups, classified according to average speed of response in word recognition. In the fast group, the mean RT was 696 ms; in the slow group the mean RT was 1,043 ms. Productive vocabulary was assessed by parental report using the MacArthur Communicative Development Inventory.

type: infants were more accurate on whole-word trials ($M = 76.85\%$, $SD = 15.43\%$) than on partial-word trials ($M = 69.35\%$, $SD = 21.49$), $F(1, 59) = 5.48$, $p < .05$. The relations of response speed and vocabulary size to accuracy in recognizing whole and partial words are illustrated in Figure 4.

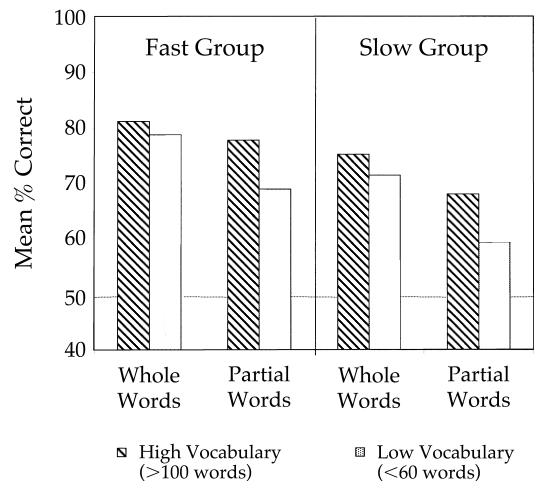


Figure 4 Accuracy on whole- and partial-word trials in relation to response speed and vocabulary size. The mean percentage of correct responses is shown for infants grouped by MacArthur Communicative Development Inventory scores (high versus low vocabulary) and by average reaction time (fast versus slow response speed). In the fast group, the mean RT was 696 ms; in the slow group the mean RT was 1,043 ms. This analysis included all trials on which infants initially fixated on either the target or the distracter picture.

GENERAL DISCUSSION

Adults can recognize a spoken word rapidly on the basis of initial acoustic–phonetic information sufficient to distinguish the word from other alternatives. The first major finding in the present research was that 18- and 21-month-old infants were also able to make use of partial information in spoken word recognition. Infants at both ages associated the auditory stimulus with the appropriate picture after hearing approximately the first two phonemes of the word. Given previous research indicating that school-aged children are slower and less accurate than adults in identifying spoken words based on partial information (Walley, 1993), it was unexpected that children as young as 18 months of age would be able to identify words from fragments. These results extend the findings of Fernald et al. (1998) and Swingley et al. (1999), which suggested that 2-year-olds process familiar words incrementally, responding before the acoustic offset of the word. In those studies, however, infants always heard the entire word so it was not possible to determine exactly how much of the word they had processed before initiating a shift in gaze. Because word fragments were used in the present study, infants heard no more than 300 ms of the target word on partial-word trials. Correct responses in this situation demonstrated unambiguously that word-initial phonetic information was sufficient to enable recognition of the word.

Although it is impressive that infants just learning to speak can identify words when hearing only brief fragments, the ability to process speech incrementally may be quite rudimentary at this age. The task used with these infants was considerably less demanding than methods used to study lexical access with older subjects. In a typical gating or lexical decision experiment with adults, the domain of candidate responses includes the participants' entire lexicon, whereas in the present experiment it may have been limited to words activated by the two pictures shown to the infants. Although recent research using this task with 24-month-olds suggests that children do not in fact restrict their lexicon to only those words related to the displayed pictures (Swingley & Fernald, *in press*), it is conceivable that for younger infants the display does serve as a constraining context. It is not known whether infants can make use of partial phonetic information when words are presented in other contexts as well. Walley (1987) found that older children could only exploit word-initial information when words occurred in highly constrained contexts, suggesting that the incremental processing strategies used by adults continue to develop in childhood. But Walley's con-

clusion that children's word recognition strategies are fundamentally different from those of adults is not plausible given the present study's finding that infants make efficient use of partial phonetic information when matching word fragments with pictures.

The second major finding in this study was that accuracy in on-line comprehension is related to lexical development. Infants with more than 100 words in their productive vocabulary were significantly more reliable in identifying spoken words than were infants with fewer than 60 words. Early links between speech production and comprehension have been documented before, primarily through observational studies showing a correlation between the number of words spoken and understood by children in the second year (Bates et al., 1988; Harris et al., 1995). The findings of this study illuminate the relation of early productive and receptive language skills from another angle, showing that more advanced lexical development is also related to efficiency in speech processing.

When hearing sentences containing a whole or partial target word, those children with larger vocabularies were significantly more reliable in orienting toward the appropriate picture than were the children who had not yet begun the vocabulary spurt. Previous research showing an increase in response speed between 18 and 24 months (Fernald et al., 1998) led to the expectation that larger vocabulary size also might be associated with faster reaction times, as well as with greater accuracy. The initial reaction time analyses, however, were inconsistent for several reasons, including the problem of subject exclusion due to missing reaction time data on partial-word trials. Another potential factor is that for some subjects there were too few responses to yield stable means. Although the accuracy scores reflected performance on all eight test trials in each condition, response latencies could only be calculated for those trials on which the infant initially fixated on the distracter picture and then shifted to the target picture. Because most infants initially fixated on the distracter picture about half of the time, their mean reaction times to whole and partial words were typically based on three to five responses per trial type. It was inevitable, however, that by chance some infants had all or a majority of target-initial trials in one or the other condition. In this case, no reaction time could be calculated, or the mean was based on only one or two responses on either whole- or partial-word trials. Because a single response that was particularly fast or slow had a substantial influence in such cases, variability among the individual reaction time means was increased. Given the limitations of the reaction time data and the partitioning of the sample by age, trial type, and vo-

cabulary size, it is not surprising that the pattern of results for the two age groups appeared somewhat inconsistent.

By grouping infants according to their average reaction time on whole-word trials, more convincing evidence was found for a relation between response speed, response accuracy, and productive vocabulary size. Those infants who were faster in recognizing intact familiar words were also significantly more accurate, and had significantly larger speaking vocabularies. Taken together, these findings suggest that more advanced lexical development is associated with greater efficiency in spoken language processing. One possible explanation for this relation is that developmental differences in the maturity of lexical representations account for the performance differences between infants in the high- and low-vocabulary groups. Children with larger vocabularies may be more accurate in recognizing familiar words because they have more mature, well-specified lexical representations, whereas those with smaller vocabularies may have incompletely specified lexical representations and therefore require more acoustic evidence for a word in order to recognize it. In this case infants in the low-vocabulary group would be expected to have more difficulty identifying word fragments relative to the high-vocabulary group. This is not what was found, however. Although infants in the low-vocabulary group were somewhat less accurate on partial-word trials than were infants in the high-vocabulary group, this difference was not reliable. Thus, no evidence was found that high- and low- vocabulary infants were differentially affected by the partial-word manipulation. Other recent research using similar procedures suggests that children at this age have well-specified lexical representations. Swingley and Aslin (2000) found that even infants with very small vocabularies rapidly distinguished correct and incorrect pronunciations of familiar words.

Even if performance differences between the high- and low-vocabulary groups are not attributable to representational differences, they may still be linguistically relevant. They could be related to some aspect of the linkage between the sound form of the word and its meaning. For example, the group differences between high- and low-vocabulary infants may be akin to within-subject frequency effects in adult word recognition. For reasons that are not precisely understood, high-frequency words are responded to more quickly and accurately than low-frequency words in a wide variety of tasks. Perhaps the high- and low-vocabulary infants in this study had differential experience with the words that were tested. Another form of differential experience could be related to the age of ac-

quisition of the target words used in this research. It seems likely that the infants with larger vocabularies had begun to recognize and produce these words at a younger age than the infants with fewer than 60 words and thus had had more extensive experience with the target words by the time of testing. Without further evidence on how word frequency and age of acquisition relate to accuracy of word recognition in individual infants, however, this account remains speculative.

The more advanced performance of infants in the high-vocabulary group could also reflect developmental differences in other cognitive processes that are not specifically associated with speech processing, but are required for successful performance in this task. It is important to consider what the infant must know and do in order to respond correctly in this procedure. The infant must first encode the picture, listen to the sentence with attention to the target word, and decide whether the spoken word matches the fixated picture. If the spoken word matches the picture in view, the appropriate response is to remain on that picture; however, if there is a mismatch, it is appropriate to reject the picture in view and mobilize a shift in gaze to the alternative picture. The accuracy of the response in either case may be influenced by any or all of the following capabilities: attentiveness to the task, speed of encoding of the visual images, association of the pictures with the appropriate words, integration of visual and auditory input, mobilization of oculomotor responses, and ability to disengage from one picture to attend to another. Developmental differences in infants' basic categorization abilities, as well as in their experience with particular objects, could also influence their performance—for example, differences in how readily the pictures of dogs used in this procedure were identified as exemplars of the category “dog.”

These are only some of the perceptual, motor, and cognitive activities that could contribute to the greater accuracy of the high-vocabulary infants in this word recognition procedure, and all entail processes that are not exclusively linked to language. This is not to say that these nonlinguistic capabilities are unrelated to language development more broadly; all may be involved in spoken language understanding in the child's daily life as well as in the experimental procedure used here. Bloom (1993) has argued that the vocabulary spurt is associated with (and perhaps mediated by) developmental changes in attentional, perceptual, and other cognitive capacities that enable the child to use more and different cues in recall and retrieval. This view is supported by recent research showing that infants with fewer than 50 words in their productive vocabulary were more dependent on

contextual cues and retrieved familiar words less reliably than infants who had entered the vocabulary spurt (Dapretto & Bjork, 2000). The point to be made is that speed and accuracy in word recognition may improve as vocabulary grows, not only because the child's lexical representations are becoming more stable but also because many other cognitive processes supporting spoken word recognition are developing at the same time.

Keeping in mind the range of cognitive factors that may influence infants' speech processing efficiency, the results of this study reveal a link between lexical development and accuracy in spoken word recognition, although the nature and direction of this relation are unclear from these findings. It could be that growth in speech processing abilities precedes and enables more rapid learning of new words, just as development in categorization abilities appears to precede the onset of the vocabulary burst (Gopnik & Meltzoff, 1987); or it could work the other way around: infants with larger lexicons may require more refined and efficient word recognition skills to distinguish among greater numbers of potentially confusable representations in the mental lexicon. Charles-Luce and Luce (1990) and Walley (1993) have argued along these lines, suggesting that increases in the size of the child's lexicon lead to the development of more efficient processing strategies and new forms of phonological organization (see Stager & Werker, 1997). Although no evidence was found that children's representations of words are underspecified or deficient, a full understanding of how growth in speech processing abilities is related to growth of the lexicon will require longitudinal research with infants across the second year.

Whatever the direction of influence, it can be concluded that when the rate of vocabulary acquisition begins to increase toward the end of the second year, infants also develop greater facility in understanding familiar words in fluent speech. Being able to identify a word on the fly has obvious benefits—for example, cognitive resources can then be allocated to understanding the words that follow. Rapid recognition of words as speech unfolds frees attentional resources for other tasks involved in understanding language. According to many models of adult sentence processing, a major constraint on comprehension is whether the listener has sufficient working memory to retain parts of the sentence that have not yet been interpreted (Caplan & Waters, 1999; Just & Carpenter, 1992). For the young language learner, it may be particularly difficult to keep track of incoming speech sounds. If, however, familiar words can be identified rapidly and accurately on the basis of the minimum

phonetic information necessary for identification, this capacity could help infants to remember what has come before, attend to novel aspects of the sentence, and parse and comprehend longer and more complex strings of words. Because figuring out the rules of syntax requires comprehending relations among non-adjacent linguistic elements across the clause, efficient identification of words in continuous speech is essential to further development in this domain. The emerging ability of infants as young as 18 months to seek out the appropriate visual referent after hearing only half a word is evidence that this efficiency begins to develop in the earliest stages of building a lexicon.

ACKNOWLEDGMENTS

This research was supported by grants from the National Institutes of Health (MH41511) and from the Office of Technology Licensing at Stanford University. The authors are grateful to Heather Birks, Amy Perfors, Kalee Magnani, and Tomoko Wakabayashi for their many contributions to this study and to Dan Lee for software development. Thanks also to the research assistants at the Center for Infant Studies at Stanford for their skillful help and dedication, and to the parents and infants who participated in this study.

ADDRESSES AND AFFILIATIONS

Corresponding author: Anne Fernald, Department of Psychology, Stanford University, Stanford, CA 94305; e-mail: fernald@psych.stanford.edu. Daniel Swingley is at the Max-Planck-Institute für Psycholinguistik, Nijmegen, The Netherlands. At the time of this study, John P. Pinto was also at Stanford University; he is now at Interval Research Corporation, Palo Alto, CA.

REFERENCES

- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.
- Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge, U.K.: Cambridge University Press.
- Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language learning*. Hillsdale, NJ: Erlbaum.
- Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression*. Cambridge, U.K.: Cambridge University Press.

- Canfield, R. L., Smith, E. G., Brezsnayak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development, 39*(2, Serial No. 158).
- Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77–126.
- Charles-Luce, J., & Luce, P. A. (1990). Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language, 17*, 205–215.
- 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., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.
- Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32*, 391–210.
- Dapretto, M., & Bjork, E. L. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development, 71*, 635–648.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethicke, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).
- Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science, 9*, 228–231.
- Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development, 58*, 1523–1531.
- Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics, 38*, 299–310.
- Haith, M. M., Wentworth, N., & Canfield, R. L. (1993). The formation of expectations in early infancy. *Advances in Infancy Research, 8*, 251–297.
- Harris, M., Yeles, C., Chasin, J., & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension and production. *Journal of Child Language, 22*, 1–18.
- Hood, B. M., & Atkinson, J. (1993). Disengaging visual attention in the infant and adult. *Infant Behavior and Development, 16*, 405–422.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102.
- Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 576–585.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1–86.
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34*, 191–243.
- Sound Edit 16 (Version 2) [Computer software]. (1996). San Francisco, CA: Macromedia.
- Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature, 388*, 381–382.
- Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76*, 147–166.
- Swingley, D., & Fernald, A. (in press). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language*.
- Swingley, D., Pinto, J. P., & Fernald, A. (1998). Assessing the speed and accuracy of word recognition in infants. In C. Rovee-Collier, L. P. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12, pp. 257–277). Stamford, CT: Ablex.
- Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*, 73–108.
- Walley, A. C. (1987). Young children's detections of word-initial and word-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 development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*, 286–350.